

DOCUMENT RESUME

ED 106 152

SE 019 187

TITLE Council of Europe Information Bulletin 1/1975.
INSTITUTION Council of Europe, Strasbourg (France). Documentation Center for Education in Europe.
PUB DATE Mar 75
NOTE 83p.
EDRS PRICE MF-\$0.76 HC-\$4.43 PLUS POSTAGE
DESCRIPTORS Bulletins; *Comparative Education; Curriculum Development; *Educational Change; *Educational Research; Evaluation; *Pilot Projects; *Program Evaluation
IDENTIFIERS *Council of Europe

ABSTRACT

This bulletin provides transcripts of lectures, reports of discussion groups, and background documents from the Educational Research Symposium on the Evaluation of School Reform Pilot Projects held at Rheinhardswaldschule in October, 1974. The symposium was organized by German authorities under the auspices of the Council of Europe and was attended by delegates from nineteen countries. The main focus of the symposium was the general trend toward school-based, decentralized innovation. Lectures presented dealt with comparative aspects of evaluation, decentralized curriculum development as action research, the relationships between school administration and evaluation of pilot projects,, problems of evaluating decentralized vs. centralized innovation, local variation and their implications for evaluation, teacher behavior as a variable in evaluating reform, and evaluation as a system component of educational development. A summing-up lecture and reports from three discussion groups are included, as are six background documents, one from Austria, four from the Federal Republic of Germany, and one from Norway. (SD)

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

COUNCIL OF EUROPE

INFORMATION BULLETIN

documentation centre for education in europe

ED106152

SE 019 187

March 1975

Contents

The evaluation of school reform pilot projects	page
Lectures	
Comparative aspects of evaluation of school reform pilot projects, W. Mitter	4
Decentralised curriculum development in the form of action research, W. Klafki	13
School administration and evaluation of pilot projects, E. Voigt	22
From centralised to decentralised innovation — Problems of evaluation, O. Lindal	27
Local variations in national curriculum development projects: self-generated school-based projects and their implications for evaluation, F. H. Sparrow	32
Teachers' behaviour as a variable in evaluating school reform pilot projects, L. Legrand	37
Evaluation as a system component of educational development, G. Petri	41
Summing-up lecture, W. Mitter	47
Reports of the Discussion Groups	
Group discussion I and II, B. Gaarder	49
Group discussion III, B. W. Kay	50
Group discussion IV, M. Gauthier	53
Background Documents	
Austria, G. Petri	55
Federal Republic of Germany — Selected projects of evaluation of school reform	63
Bavaria, D. Krywalski	63
North-Rhine Westphalia, E. Parey	66
Ruhr Teachers' College, H.-G. Rolff	67
Baden-Wuerttemberg, R. H. Weiss	69
Norway, O. Lindal	72

The Information Bulletin which is published free of charge in an English and a French edition reprints important policy documents of European interest in educational, cultural and scientific fields.

IMPORTANT NOTICE TO READERS

Readers are asked to note that, from now onwards, as part of a general reform of the information and documentation services relating to the educational and cultural work of the Council of Europe, this Bulletin will no longer contain a First Part summarising the results of European meetings held under the auspices of the Council for Cultural Co-operation. Such information will be inserted, in an abbreviated form, in the review "Education and Culture" which is already received by most readers of the Bulletin.

Two further changes should be noted. First, the Bulletin will be published in three, four or five issues per annum, according to the nature and the volume of the material to be presented. Second, the scientific or policy papers resulting from European conferences, colloquies or symposia, which henceforth take up the whole of the Bulletin, will cover the problems of cultural development as well as educational themes.

Specialists who are particularly interested in *national* educational reforms in Europe are reminded that abstracts describing such reforms continue to be published in the "Newsletter" of the Documentation Centre for Education in Europe.

THE EVALUATION OF SCHOOL REFORM PILOT PROJECTS

Educational Research Symposium, Rheinhardswaldschule, 14th—18th October 1974

The Educational Research Symposium on the Evaluation of School Reform Pilot Projects was the sixth in a series of symposia sponsored by the Educational Research Committee. The Symposium was organised by the German authorities under the auspices of the Council of Europe; and was held at Rheinhardswaldschule near Kassel from 14th to 18th October 1974. The Chairman was Professor W. Mitter, Director of the German Institute for International Educational Research, Frankfurt/Main. Delegates from nineteen member States — researchers, officials from education ministries and representatives of school administrations — participated in the Symposium.

The objective of the Symposium was to compare methods and results of the evaluation of school reform pilot projects and to investigate the scope for European co-operation in this field. The evaluation of reform has become one of the most vital issues along the borderline between policy and research. The rather naïve judgments on innovation which are, in the battlefield of ideologies, based on an act of faith, are more and more being superseded by scientific methods of evaluation of change. It was acknowledged, however, that differing objectives in school reform made the universal application of specific evaluation methods doubtful and that evaluation by researchers could also be abused for political ends.

The Symposium dealt in particular with the general trend towards school-based, de-centralised innovation. The centre-periphery model in which reform is conceived at the Ministry and imposed hierarchically on the schools and their population has in many countries not been able to live up to the expectations which backed this model. A new model, the periphery-centre-periphery model, is therefore developing in a number of countries. At present, innovation often begins at the grassroots and takes the shape of pilot projects in which teachers and administrators, parents and pupils, and even the whole local community, take an active part. Such innovation is then reported to the Ministry by the Inspectorate, by written reports or by demonstration of its results. It finally spreads from the centre back to the periphery in a general dissemination process which may be strongly reinforced if the evaluation can claim to be research-based.

Thus the Symposium provided a useful opportunity for comparing developments, results and methods in the various countries. General agreement was reached that the most useful evaluation projects are those in which the research teams are not neutral observers applying tests at periodic intervals from the "outside", but those projects where the researchers are active participants, together with teachers, administrators and parents, from the very beginning of the evaluation to its end.

*All lectures, the discussion group reports and the background documents dealing with national evaluation projects are given below. A German version of the Symposium papers will shortly appear in a special issue of *Bildung und Erziehung*.*

LECTURES

Comparative aspects of evaluation of school reform pilot projects

by Professor W. MITTER, Director,
Deutsches Institut für Internationale Pädagogische
Forschung, Frankfurt/Main.

Introduction

School reform pilot projects are today being carried out and evaluated in all industrial countries. Thus the Council for Cultural Co-operation of the Council of Europe deserves thanks for making this symposium possible, affording educational scientists and representatives of school administrations from the member countries of the Council of Europe the opportunity to discuss the evaluation of school reform pilot projects from the point of view of general and specialised theoretical research, their role in the formulation of educational and social policy, and the role of the educational scientist in the research process, and to compare experience regarding the planning, conceptualisation and execution of school reform pilot projects in general and selected projects in particular.

The programme of lectures is concentrated on the decentralisation of school reforms, the development of school-based evaluative projects and illustration of the interaction between aspects of organisational theory and methodology. The lecture titles themselves, with their high level of abstraction, show that the subjects selected do not directly relate to national experience. This is deliberate, insofar as any tendency to reduce the programme to an additive sequence of national reports was to be avoided. It is true that an analysis of national peculiarities would have given the participants items of valuable information, but it would have made the effort to define general methodological questions as well as the attempt to find and develop comparison criteria more difficult. For national reports — especially when they have the character of a survey — tend toward a global view of the subject, even when the latter is relatively limited, as in our case it is limited to school reform pilot projects.

At the same time, the decision of the preparatory committee to define the lecture subjects in terms of theoretical research categories does not ignore the fact that what is said by the lecturers from various countries will reflect experience gained in concrete project work carried out under conditions set by norms of national educational policy and

scientific tradition. The supra-national significance of the entire subject field, as well as the circle of participants, will also help in explicitly drawing international comparisons in group work and — even more important — in guiding the discussions implicitly in the sense of a "hidden curriculum". This will happen despite the fact that international comparison of concrete evaluation projects is the specific subject neither of the entire symposium nor of individual lectures. The method of beginning with the problems expounded in the lectures and with the abstraction of experience based on analysis of national specifics, and of stimulating the participants to make explicit and implicit international comparisons, appears more promising. Perhaps the course of this symposium will confirm this assumption — or perhaps it will disprove it and thus indicate alternative approaches for future symposia.

I should like, by way of introduction to our discussion, to draw your attention to some of the comparative aspects of the evaluation of school reform pilot projects. In the first part of my remarks I would like to discuss arguments in favour of legitimising international comparative studies within educational research as a whole. The second part is devoted to an outline analysis of controversial ideas put forward in the discussion of evaluation theory in the United States during the 1960s and in recent discussion of the function of comparison within evaluative research in the Federal Republic of Germany. The insights gained from both points will form the basis of an attempt to outline — in the conclusion — opportunities for and limitations on international comparative studies which could relate specifically to school reform pilot projects.

International comparative studies in educational research

1. What arguments could be given in order to justify the necessity of international comparative studies in educational research?

1.1 As an instrument of the search for truth and hence the humanisation of man's society, science fulfils the task of critically testing and clarifying naïve, pre-scientific descriptions and explanations of conditions and events in organic and inorganic nature — and thus also in the field of social interaction — wherever these are found. Comparative international studies of situations and events in the field of education conform to this definition of the task, for pre-scientific international comparisons have been made ever since there have been rudimentary forms of that social entity, the "nation". Just as the rules of education in African bush schools were altered as a result of comparison with analogous precepts and rites discovered and observed in neighbouring tribes, so did important 19th century representatives of the intellectual life of their countries travel abroad to visit schools and observe lessons there. From their observations and comparisons, they formed recommendations aimed at showing up foreign educational practice as an example to be followed or rejected. The great pedagogical travellers of the 19th century, such as Victor Cousin, Matthew Arnold, Horace Mann and Leo Tolstoi, are honoured in comparative education circles as forerunners and pioneers of the 20th century educational studies.

We realise, of course, that pre-scientific comparisons can lead away from scientific understanding of reality and even trigger irrational processes in opinion forming. The epistemological function of comparative education research is emphasised by the presence of such processes. In the Federal Republic of Germany, for instance, the term "Swedish school reform" has been degraded to a provocative slogan as a result of naïve comparisons by both supporters and opponents of the comprehensive "Gesamtschule"; the same thing happened to events and measures which became known in fragmentary and distorted fashion from the American educational system. The fact that naïve comparison is also capable of producing paradoxes is shown by the reception given to information about the educational system of the German Democratic Republic; the paradox here is that approval is based not only on ideological commitments to socialist quality (not under discussion here), but also on superficial admiration of "discipline" and "achievement" which betrays a conservative attitude and does not take into account the political, social and ideological background to individual educational and pedagogical measures.

1.2 Comparative international studies are prompted by the logical discovery through research that there are phenomena in the field of education which can

be explained *only* in terms of national circumstances or which are *best* explained with reference to the nation as a social entity. Hence the choice of the nation as a unit of study for certain topics of educational research. For most of Europe this at least seems to provide a working definition, as is very well evidenced by the contents of the survey on educational research policy in Council of Europe member states, produced in 1973 by the Documentation Centre for Education in Europe. While contenting ourselves with a working definition, we by no means fail to realise that thorough clarification of concepts constantly requires new theoretical efforts. These need to be tested empirically as well as backed up by reference to the historical development of the concept of the "nation", whose ambivalence is reflected in the distinction between J. G. Herder's philosophical interpretation in terms of history and culture (the "nation" as a natural and organic linguistic community) and J. J. Rousseau's emphasis on political choice (the nation as a political community determined by "volonté générale").

Comparative educational research is forced to participate in this continuous clarification of concepts by current events as well, whether these concern the federalisation of educational systems on ethnic principles, as in Belgium and Czechoslovakia, or the new minority problem of migrant workers (Gastarbeiter). These are marginal phenomena in European educational terms, important though they may be to the countries concerned. But the necessity mentioned above of using the social entity "nation" as a frame of reference to explain "normal" features of education and educational policy can be demonstrated using numerous examples. We shall select a few, first of all from the area of school structures.

Structural features of educational systems

Example 1: As soon as the topical question of access to higher education is broached in international terms, educational policy-makers and researchers are confronted with national regulations governing the relationship between school-leaving examinations and qualifications for university entrance, which raise a host of complex issues in each individual country. The organisation of a "big lift" for West German students to study at American colleges would be thus not exclusively a problem of financial policy.

Example 2: A comparative observation of the development of comprehensive schools and "Gesamtschulen" in Great Britain and the Federal

Republic of Germany must take into account the highly developed and socially accepted British system of "public" (private) schools, which offers parents who do not wish to send their children to comprehensive schools a real alternative, even when there are no state grammar schools available. In the Länder of the Federal Republic of Germany political decisions to make the integrated Gesamtschule the norm would have much more impact on the parents of individual ten or twelve-year-olds.

Example 3: The United States Supreme Court's most recent decision against the "bussing" introduced in the 1960s and against the related organisation of supra-local school administrations was based specifically on the grounds that local control of the schools was a central principle of educational policy which must not be infringed. Participants in a symposium on the decentralisation of school reforms in Europe might find food for thought in this argument, regardless of the intra-national significance and effect of this ruling.

These three examples were chosen to demonstrate structural features of educational systems which are due to national characteristics, but they may also serve to draw attention to curriculum characteristics. Thus the relationship between the completion of secondary education and university entrance also depends on differences in curricula at equivalent levels in different countries; it should also be mentioned that the "bussing" decision marked the end of experiments in the area of social science curricula which had been supported by the State for a decade.

The function of history teaching in the process of political socialisation may be taken as a typical example of a concentrated study of curriculum characteristics; the International School Book Institute in Braunschweig has been engaged in such a study for more than two decades. The teaching of mathematics should also be analysed in the context of a comparative study of national educational strategies before interpreting the significance of international achievement tests such as the IEA's, for the criteria and test procedures used as a basis for assessing the effectiveness of mathematics teaching are fundamentally open to question. Were achievements in mathematics — we would have to ask — tested primarily from the point of view of applicability to scientific and technological disciplines, or was sufficient consideration given to the fact that a national educational system might regard it as the main function of mathematics instruction to develop young people's ability to think logically and thus afford them a basis for philosophical understanding of themselves and of the world? International achievement comparisons are prone to be misinterpreted and reduced to

mere pegs for ideological arguments if insufficient allowance is made for national background variables.

Lastly, situations and events worthy of study in national educational systems can be affected when nations are split or absorbed into new social entities as a result of political changes. In the field of education, traditional structures, contents and attitudes relating to the "old" nation outlive such changes, which can lead in comparative international studies to discoveries as theoretically intriguing as they are politically significant. One example of this is the comparative study of present-day school development in the states which emerged from the Austro-Hungarian monarchy, although here the question of whether the term "nation" can be applied to nationality states is a matter for further inquiry and reflection.

1.3 In the age of international communication, the development of clear *terminology* is of the greatest importance, as may be seen for instance from the compilation of multi-language thesauri. As regards the contribution comparative international studies can make in this respect, two types of terminology deserve special mention.

Where research theory and method are concerned there is a supra-national language commonly used in the empirical social sciences. As regards the subject of this symposium, for example, I am thinking of terms used to define research models, research types, investigative procedures and data-gathering methods. Linguistically of course, this supra-national language is derived primarily from one national ie, English, source. This concentration makes it easier to acquire and use sociological vocabulary, not only for native speakers of English, but also for "foreigners", although misunderstandings may occur owing to the use of English terms in other languages without due regard for semantic factors. Accordingly, in formulating the overall subject of this symposium, it had to be taken into account that the German term "wissenschaftliche Begleitung" has no literal counterpart in English or French and that the English term "evaluation" is not covered by the German borrowed term "Evaluation". The danger of misinterpretation is specially large when theoretical ideas and findings are communicated to school administrators and teachers concerned with putting them into practice. The involvement of "practitioners" in actual research, as in action research, increases this danger but could, due to the permanent necessity of clarifying ideas in practical everyday work, help to reduce the number of "fashionable" terms and thus — in the field which concerns us here — to develop meaningful supra-national scientific language.

Much more difficult than fixing research theory and method terminology is the attainment of linguistic clarity in the description of national educational systems within which — in our case — the school reform pilot projects and evaluation projects to be analysed are carried out. This applies not only to the structural framework, but also to curriculum development and to socio-pedagogical factors in the broadest sense of the word. Major problems arise here not only in the dissemination of research findings, but in the course of research itself, at the definition stage. The analysis of general aims and the definition of operational objectives, to take one important example, require a consensus on the terms to be used *before* content factors are worked out unless the discussion is to be reduced to mere controversy or refuge is to be sought in terms which are only apparently unambiguous. An example is the French term "démocratisation", which refers to equal opportunity in school attendance and achievement, and does not mean the same thing as the West German "Demokratisierung", which expressly covers the extension of pupils', teachers', and parents' participation in decision-making. Misunderstandings are particularly frequent in the field of political education, where national usage is overlaid with non-specific ideologically slanted terminology; the confusion is particularly starkly revealed in the unthinking use of such basic concepts as "democracy" and "peace".

The classification of school structures is complicated by the fact that the terminology tends to reflect historical tradition rather than present-day functions. Traditional terms remain in use and are even — as in the case of the "Gymnasium" in Czechoslovakia — re-introduced, even where the types of school to which they refer have changed radically in the meantime in both structure and curriculum. One example may suffice to illustrate the problem: in the present school system of the German Democratic Republic the term "Oberschule" denotes classes 1 to 12 of general education, which as the "Allgemeinbildende polytechnische Oberschule" is the general school for classes 1 to 10 for all schoolchildren and has nothing in common with the former selective German "Oberschule". Problems of understanding which arise here simply from a "German-with-German" comparison are further increased if, for example, the American "high school" is translated into German without qualification as "Oberschule".

Pragmatic considerations

1.4 Whereas the arguments advanced so far justify the choice of the nation as a unit of study on

theoretical grounds, the remaining two are based on pragmatic considerations. The first is that in gathering data educational research can most readily draw on the data bases prepared by the documentation centres of individual countries and supra-national institutions — such as the Documentation Centre for Education in Europe and the International Bureau of Education, which are continuously kept up to date and organised, the supra-national centres basing their own classification on national categories. This affords educational research access to material which could otherwise be obtained only with great difficulty and subject to limitations or not at all. Suppliers of data are, on the one hand, the central statistical offices (with their branches) which compile educational and social statistics, and on the other the documentation services, which systematically collect primary verbal sources such as laws, regulations, statutes, etc.

The abundance of data thus classified encourages comparisons, and the increasing internationalisation of data collection creates new opportunities. However, there are dangers for the educational scientist who seeks to discover pedagogical and anthropological phenomena, as may be illustrated by two problems:

Problem 1: The statistical data bases and verbal documentation prepared by state agencies are first of all incomplete and — with regard to international applicability — are not compiled according to standard criteria. Although considerable improvements and refinements may be expected here in the near future, the more central state bodies monopolise data collection and storage, the less this is true of doubts about the reliability of published data. Comparative studies which go beyond the context of the national state as a self-contained unit and select regional phenomena such as evaluative research projects, for instance, as subjects of comparison, would be welcome from the point of view of complementary data collection and hence for the questioning and correcting of official national statistics alone.

Problem 2: Another problem is that the relatively easy access to data tends to make educational science dependent on the "pedagogy of the school" and reinforces a tendency to over-estimate the influence of public and private institutions on the young and even on adults. One does not have to be a "de-schooler" to have doubts about such tendencies. The distinction which has become current in the empirical social sciences between "international" and "inter-cultural" comparison, concentrating on comparing socialisation processes in social groups, is certainly correct and necessary in cate-

gorical terms. If, however, this should lead to the creation of distinct types of research pursued independently of each other, educational theory and practice would both suffer in that those involved in the educational process — teachers and pupils — would be investigated through two isolated comparison procedures: on the one hand through de-politicised “inter-cultural” comparison and on the other through an “international” comparison in which the working concept “nation” became a stable term whose use could lead to exaggeration of the importance of national background variables.

1.5 The last argument can be stated very briefly: comparative international studies are prompted and justified by the increasing *internationalisation of educational policy* and by the world-wide and regional activities of international and supranational organisations.

2. We now turn to the second question, the problem of the *applicability and usefulness of comparison* in evaluative research.

Outline analysis of controversial ideas in the evaluation theory

2.1 Let us begin with opinions on comparison as a research technique that were defended in the 1960s in theoretical discussion in the United States on the question of curriculum evaluation. The controversy was directly triggered by the question of whether new curricula should be evaluated by comparison with control groups taught according to conventional curricula. It had been found that efficiency tests frequently showed no significant differences. The question was accordingly: can or should new curricula be subjected to evaluative comparison with conventional curricula? The two conflicting answers to this question were represented in the American literature by Lee J. Cronbach and Michael Scriven. The essential arguments in this controversy are outlined below. It is emphasised that this brief analysis does no more than describe examples, and in particular that the later research findings of the writers mentioned are left out of account. Cronbach (1963) contests the value of comparative evaluation. Firstly, he draws attention to its inconclusiveness, as revealed in the non-significant differences in average test results, and to sources of error attributable to the teacher's commitment, which can scarcely be measured and which in turn influences the pupils' motivation during the experiment. It is never possible to say with certainty which variables explain the observed and tested advantages of a new curriculum. Even

in the case of extensive, well-planned and correctly controlled comparative studies therefore, it is dangerous to generalise from the results. Secondly, Cronbach raises the question of the often different objectives of the curricula which are being compared. Except in the rare cases where one of the curricula could be clearly shown to be better, comparative interpretation is hampered by the difference in objectives. Finally, he raises the pragmatic question of the financial outlay in relation to the expected significance of the results; he regards this question as important enough for the usefulness of a comparative evaluation to be assessed according to whether the ratio between the two quantities is reasonable — which, as a rule, it is not.

Naïve and scientific comparison

Michael Scriven (1967) takes the opposite view that experiments with new curricula are always stimulated and spurred on by competition and that curriculum evaluation should therefore never exclude comparison; in this connection I would like to call to mind the relationship mentioned at the outset between naïve and scientific comparison. Cronbach's example of the new car model illustrates the controversy. Cronbach's contention that the engineer is interested only in the power and reliability of the model he is testing is answered by Scriven by saying that the only yardstick of a car's power and reliability and our interest in it is our knowledge of what has proved possible within a certain price range, a certain amount of room and a certain total weight.

As regards the significance of comparative findings, Scriven basically agrees with Cronbach that the instruments used for comparative evaluation are not yet sufficiently well calibrated. He concludes from this, however, that the financial outlay needed to make the tests more precise is justified, as is an expansion of comparative groups beyond the sizes customary hitherto.

Scriven touches upon the question of the applicability of comparative evaluations to educational policy to the extent that, in his opinion, the values assigned to the content and objectives of curricular innovations must be suitably weighted; here even relatively small differences in achievement might prove significant. This applies to the evaluation of both new and conventional curricula. In his view Cronbach does not fully appreciate the fact that the absence of clear differences frequently provides the very evidence needed to decide in favour of developing tried and tested curricula rather than introduce promising, but radically new

models. It might be equally wrong, however, to discontinue an experiment at an early stage because comparative evaluation in the first trial had not revealed significant differences, especially as the results would probably not provide any information about the possible long-term effects of curriculum changes.

Scriven's theoretical reflections later taken up by Robert E. Stake (1968) and supported by Richard C. Anderson (1969), extend beyond the question of instrumental precision and applicability to educational theory. Here we are concerned with the answer to Cronbach's question about differences in objectives as an obstacle to comparative evaluation. According to Scriven, in the absence of comparison curriculum evaluation could be carried out only by means of tests against absolute standards. This, he claims, raises the question of absolute criteria, especially as experience has shown that scales used in non-comparative evaluations are usually percentage scales or scales with implicit comparisons. Thus terms such as "useful" or "valuable contribution" as value judgements of curricula always contain an element which indicates their superiority over other curricula, with which they are implicitly compared.

Stake traces the setting of absolute standards to personal judgements on the part of the evaluators; it seems to me that attention should rather be drawn to the influence of economically, politically and ideologically determined group interests. With his evaluation matrix, Stake has devised an instrument for the orientation of objectives in terms of absolute as well as relative standards; it emphasises the value of evaluations which include the comparison of relative standards in determining positions on the scale.

The relevance of the research theory controversy to educational decision-making is touched upon by Stake in raising the question of target groups. He argues that the practical teacher who actually has to make a choice of curriculum will tend to find Scriven's positive attitude to comparative evaluation more plausible than the curriculum innovator or the educational technologist.

2.2. Most of the arguments expounded in this discussion of evaluation theory have been repeated in recent controversies about the use of comparative studies in evaluative research in the Federal Republic of Germany. I shall confine myself here to opinions formulated by representatives of school administrations or members of state institutes of educational research.

They refer to the recommendation passed by the Education Commission of the German Educational

Council (Bildungskommission — Deutscher Bildungsrat) on January 31 1969, on the "Introduction of Comprehensive School Experiments" which also emphasises scientific control of the projects and the significance of evaluative research.

The Educational Council's recommendation distinguishes three types of pilot-project evaluation:

1. Comparisons of the system as a whole with the traditional tripartite school system ("inter-system comparisons"),
2. Comparisons of the structures of the new system with its intentions ("system-immanent comparisons"),
3. Comparisons of variants within the comprehensive school system ("intra-system comparisons").

Before going into the various reactions to this recommendation, I should like to point out that the experimental situation differs from curriculum innovation in America by its greater breadth and complexity. The West German comprehensive school pilot projects are concerned not only with testing new curricula, but also with testing a type of school which differs so greatly from traditional schools in the tripartite system with respect to both pupil population (social origin, level of knowledge, intelligence) and objectives that the differences in objectives discussed by American educational scientists stand out much more sharply. In addition, the introduction of the new school type necessarily entails testing forms of organisation based on external and internal differentiation which unlike similar arrangements in traditional secondary schools are central to the system.

For the purpose of illustrating the main issues, we shall take opinions from Baden-Wuerttemberg, Bavaria, Hesse and Lower Saxony, concentrating on the centres of controversy and beginning with the general remark that all the authors are aware of the subtle difficulties with which project planning has to cope in actual practice.

Need for system immanent comparisons

There seems to be general agreement on the need for system-immanent comparisons covering parallel courses in the same pilot school, they are valued as an instrument for improving the results of experiments, as well as for correcting partial objectives. The authors differ, however, as to the possibility of intra-system comparisons and especially of inter-system comparisons.

Strong reservations can be found in the documentary material ("Unterlagen") published by the Ministry of Education of Lower Saxony. Doubt is

cast on the usefulness of exact comparative studies of the Gesamtschule with the traditional school system in general, on the grounds that there is no common basis of comparison owing to differing objectives and marginal conditions. The reservation about intrasystem comparisons is limited in that thematically limited projects are considered possible under certain circumstances. But even this concession is outweighed by the negative overall attitude which classifies comparisons of different Gesamtschule projects as "methodically doubtful".

The evaluative research projects initiated by the Hesse Ministry of Education point in the same direction, particularly in the reference to a gradual shift from efficiency tests and system comparisons in favour of system-immanent studies. The only reason given for this is the discrepancy between the two school systems' objectives and the means of attaining them.

The Bavarian State Institute of Educational Research and Planning holds the opposite view. The evaluation of comprehensive school pilot projects is clearly slanted towards efficiency control; inter-system comparisons are accepted and control schools (from the tripartite system) are included in the research schemes.

The strategy worked out by the Institute of Educational Planning and Information in Baden-Wuerttemberg resembles the Bavarian approach in that it provides for all three types of comparison. However, intra-system comparisons and especially inter-system comparisons are made subject to a number of criteria and conditions: guarantee of random sampling in school selection, great homogeneity of control population, multivariate character of the project plan, consideration of the "conflict-reducing" effect — eg through the participation of psychologists — limitation of problem scope in the light of the comparison-hindering influence of differences in objectives etc. These provisos amount to a general reservation which — as far as practicability is concerned — comes close to the position adopted by Hesse and Lower Saxony, though in less drastic form.

The above remarks have been intentionally confined to opinions which — although not to be regarded as official statements of educational strategy — certainly give some clues to official policy. The differing interpretations of the German Educational Council recommendation reveal fundamental differences in the intermediate-range objectives of secondary school development, which brings us back to the target-group problem mentioned by Robert E. Stake. In this particular case the fact

that the research is commissioned by the same agency as is intended primarily to benefit from it is a major factor in the appreciation and control of comparative studies.

The opinions from Hesse and Lower Saxony are based on educational policy strategies which are oriented toward introducing the "integrierte Gesamtschule" as the norm and emphasise the discrepancies between the objectives of the desired and the traditional school system by attributing general qualities such as "equal opportunity" "social integration", "individualisation of learning" and "care for the socially deprived" only to the new system. The opinions from Bavaria and Baden-Wuerttemberg, on the other hand, draw attention to strategies in which the decision for or against the general introduction of the Gesamtschule remains open and the function of evaluative research is viewed primarily not as optimising school reform pilot projects, but as advising educational policy bodies which have not yet made up their minds on the basic issue. Where the Bavarian position is concerned, the openness of the decision as an objective is limited by the research plans to the extent that the Gesamtschulen are required to prove their quality in terms of their superiority over traditional secondary schools, whilst its reverse is not expressly provided for.

Since the types of comparison illustrated by current evaluation strategies in the Federal Republic of Germany can also be found at the level of international comparison, I do not intend to make a separate appraisal of the attitudes described. We may, however, draw the obvious conclusion that decisions for or against comparison in evaluation strategy and the grounds for such decisions appear to be influenced by the expectations of the initiators coinciding with the addressees of evaluative research, even more than by the production intentions of the research team.

Applicability of school reform pilot projects

3. Our last question concerns the applicability of school reform pilot projects for purposes of international comparison. Our answer is intended as an attempt to put forward general ideas on explaining this problem and thus to provide a framework for the discussion of concrete research projects.

3.1 It is difficult to classify comparisons of school reform pilot projects in two or more countries into one of the three categories described earlier. From a formal point of view, the subject matter places them in the category of intra-system comparison if the projects are carried out at the same or similar

levels and in the same or similar types of school. Examples might be comparisons of Gesamtschulen, comprehensive schools and collèges d'enseignement secondaire. It nevertheless appears justified to classify such comparisons also as *inter-system comparisons*, because on the basis of the input-output model the input variables, the operational variables and, especially, the output variables determined by the pilot schools' objectives and intentions differ so much in each case that formation of a hypothesis should — in order to be useful — proceed from the assumption of independent "systems" as subjects of comparison. This option is lent additional legitimization in cases where pilot projects within a school type introduced as the norm — e. g. in the upper classes of the elementary school in Sweden — are to be compared with projects which are either islands in a sea of schools of another type — e. g. Gesamtschulen in Bavaria — or have been initiated in a region which has a mixture of existing types — e. g. comprehensive schools in Great Britain or Gesamtschulen in Hesse. In order not to limit the discussion to Gesamtschule projects, the pre-school field may be mentioned as a second example. In this case pilot projects may be treated in terms of different "systems", if there are substantial differences in the proportions of four or five-year-olds in pre-school education in relation to the total populations in that age group.

The most obvious subjects of comparison are individual school reform pilot projects, although the — admittedly complicated — comparison of project planning and the progress of experiments within comparable regions should also be considered.

3.2 Inter-system comparisons on an international scale belong in the *problem approach* category as developed by comparative educational science, which has displaced the formerly dominant total analysis owing to the growing realisation that demands made by this form of study are usually too wide and cannot be satisfied by empirical research. In terms of our subject, this means that the implementation of school reform pilot projects can be defined as a "problem" and hence as a basis for comparison. The aim of the comparison could then be to establish whether and to what extent the educational policy strategies of the "home countries" are reflected in the pilot projects under comparison — in the sense of functional equivalents. A prerequisite for such project planning would be previous system-imminent comparisons within the school systems to be compared.

The purpose of inter-system comparisons on an international scale may generally be explained by reference to the arguments advanced at the outset to justify comparative international studies, in particular with the qualification of the central

argument to the effect that the existence of the nation as a social entity is most important for the preparation of a catalogue of variables — through the inclusion of appropriate background variables on the output side. This produces effects, firstly on reservations about hasty and unfiltered adaptations of foreign reform projects and, secondly, on the more precise definition of intranational evaluation projects through greater regard for national educational traditions whose suppression is always reflected post facto in the failure of such projects.

3.3 International inter-system comparisons confront a research team — if only because of the difficulties in getting started and in apportioning costs — with the necessity of planning their projects particularly carefully. This applies first of all to investigation of the *quality criteria* which have to be observed in all empirical studies. International studies can thus contribute to improving methods and instruments which are important to evaluative research in general. A quality criterion which demands additional attention and which would have to be included in a "methodology of comparison of school reform pilot projects" would be terminological uniformity, which would require fixation of a standardised language for the countries to be compared.

3.4 The successful implementation of comparative international studies would seem likely to be a way of reducing apodictic judgements on the dubious value of inter-system and intra-system comparison as research methods. This assumption refers to the question of differing objectives which, as an obstacle to the establishment of a basis for comparison has been a cause of concern to American curriculum evaluators, as well as to project planners in the Länder of the Federal Republic of Germany. However, even when the reservations resulting from a consideration of this problem have been taken fully into account, the following arguments speak strongly in favour of comparisons in the evaluation of school reform pilot projects.

First of all, I would draw attention to the significance — mentioned by Scriven and Stake — of the comparison of *standards* as socio-political and philosophical aims which determine the operational objectives of the curricula under comparison. Since in pluralistic societies educational policy-makers are confronted with the relationship between absolute and relative standards in all their decisions, it is important for educational research not to dodge this issue by alleging, insoluble methodological difficulties and relying on naïve comparison.

Secondly, the argument that the objectives of differently structured school systems cannot be compared ignore the existence of similar, if not identical, *partial objectives* which are worthy of comparative analysis, not only from the point of view of research theory but also from that of educational and financial policy. This statement could be verified for the acquisition of cultural techniques as well as for the learning of foreign languages or the investigation of topics in the natural sciences. It must also be remembered that the identity or difference of partial objectives may conflict with differences in the field of general educational aims. For instance, the debate on the modernisation of mathematics teaching, which is a possible subject for international pilot project comparison, runs largely parallel to general developments in the primary and secondary school field.

Finally, international comparisons can promote objectivity in the standards controversy, in that detached contemplation of educational systems in whose reform the educational policy-maker or researcher is not directly involved may help bring about recognition and acceptance of differently formulated standards as variations on the same absolute principles.

3.5 This detachment could also aid in clarifying the political frame of reference in which not only practical research decisions are made, but also research intentions are theoretically formed. In all open societies, it draws the attention to controversial educational policy strategies, and the educational scientist who undertakes commissioned research — in our case the evaluation of school reform pilot projects — accepts specific conditional factors. In the countries where we work, these are legitimate phenomena of educational reality. Thus there is no reason to suppress them from our consciousness, because it is the identification of them that makes it possible to subject project plans and research findings to critical scrutiny.

BIBLIOGRAPHY

BERSTECHER, D.: *Zu Theorie und Technik des internationalen Vergleichs*, Stuttgart, Klett 1970.

Council of Europe: *Educational Research Policy in European Countries. Survey 1973*, Documentation Centre for Education in Europe, Strasbourg 1974.

Deutscher Bildungsrat: *Empfehlungen der Bildungskommission*

— Einrichtung von Schulversuchen mit Ganztagschülern, 1968.

- Einrichtung von Schulversuchen mit Gesamtschülern, 1969.
- Für Förderung praxisnaher Curriculumentwicklung, 1974.
- Aspekte für Planung der Bildungsforschung, 1974.

DIECKMANN, B.: *Zur Strategie des systematischen internationalen Vergleichs*, Stuttgart, Klett 1970.

FROMMELT, B. and RUTZ, G.: „Projekte praxisbegleitender Forschung an hessischen Gesamtschulen“. In *Bildungspolitische Informationen*. Hrsg. vom Hessischen Kultusminister, Nr. 3/73, p. 11 ff.

KASTNER, R.: „Die wissenschaftliche Begleitung der Versuche mit Gesamtschulen“. In *Der Niedersächsische Kultusminister, Gesamtschule in Niedersachsen, Unterlagen zur Entwicklung von Gesamtschulen in Niedersachsen*. Hrsg. vom Niedersächsischen Kultusminister, Hannover 1974 (Januar), p. 36 f.

KRYWALSKI, D.: *Die wissenschaftliche Begleituntersuchung von Schulversuchen mit Gesamt- und Ganztagschulen in Bayern*. Unpublished working paper of the project-group „Zur wissenschaftlichen Begleitung von Schulversuchen“, Staatsinstitut für Bildungsforschung und -planung. München, 16. 7. 1974.

LANDSHEERE, G. de.: *Einführung in die pädagogische Forschung*. German Translation, Weinheim, Beltz 1969. Introduction à la recherche en éducation, 1964, 3rd revised and enlarged edition. Paris, Armand Colin-Bourrelle 1970.

ROBINSON, S. B.: „Erziehungswissenschaft: Vergleichende Erziehungswissenschaft“. In *Handbuch pädagogischer Grundbegriffe*, Band 1. München 1970.

STRAKA, G. A.: *Forschungsstrategien zur Evaluation von Schulversuchen. Probleme empirischer Begleitung und empirischer Kontrolle von Modellschulen*. Weinheim, Beltz 1974.

WEISS, R. H.: „Wissenschaftliche Begleitung der Modellschulen Baden-Württembergs. Konzeption des Instituts für Bildungsplanung und Studieninformation Stuttgart. Erste Erfahrungen — Materialien.“ Offprint from: *Die Schulwarte*, Monatsschrift für Unterricht und Erziehung, Nr. 7/1972.

WULF, CH. (Ed): *Evaluation, Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. Texte*. Piper, München 1972. Contains the quoted articles by CRONBACH L. J.: „Evaluation for course improvement“. In

HEATH, R. W. (Ed): *New Curricula*, New York, Harper & Row 1964. SCRIVEN, M.: *The methodology of evaluation. Perspectives on Curriculum Evaluation*. American Educational Research Association, Monograph Series on Curriculum Evaluation No 1. Chicago, Rand McNally 1967. STAKE, R. E.: *The countenance of educational*

evaluation. Teachers College Record 1967/68. ANDERSON, R. C.: "A comparative field experiment. An illustration from high school biology". In HASTINGS, TH. (Ed), *Proceedings of the 1968 invitational conference on testing problems*. Princeton, New Jersey, Educational Testing Service 1969.

Decentralised curriculum development in the form of action research

by Professor W. KLAFKI,
Marburg University.

Introduction

In my lecture, I should like to discuss some problems of a relatively new type of school innovation research; it is a new type, only the very beginnings of which have been realised in the Federal Republic of Germany and — as far as I know — in other European countries. Thus it remains primarily a programme up to the present. I refer to this new research as "Decentralised Curriculum Development (or school-based Curriculum-development) in the Form of Action Research". At present, this cannot be the presentation of a completely planned, detailed model; rather, I can only sketch a few basic problems, without any claim to completeness.

The basis of my remarks are on the one hand, corresponding scientific publications. On the other, I refer to perceptions and experience from two practical curriculum-development projects in which I have participated or am participating:

- a) The first was a fairly extensive attempt to revise past curricula for schools of general education in the land of Hesse, a project which ran from 1969 to 1971 and which was unfortunately interrupted by the Hessian Ministry of Culture at a time when we were just going to take the step from basic preliminary reflections to systematic combination of supraregional curriculum planning with decentralised curriculum development. (1)

- b) The second practical source of my considerations are experiences from two projects, one of which being carried out by Marburg University in cooperation with groups of teachers at the primary school and directed by myself. This is a contribution to curriculum reform in the Grundschule in the form of action research (2). The second is an action research-project too, which so far was carried out by a combined working group "Curriculum/Social sciences" at the Bildungstechnologisches Zentrum Wiesbaden and to which I was attached as a scientific advisor. (3)

In the first section, I will begin with some remarks on the main terms which I will be using.

In the new research type "decentralised curriculum development in the form of action research" — and here I come to the second point of my paper — two activities are combined, although they do not

(1) Cf. Klafki, W., Lingelbach K. Ch. and Nicklas, H. W. (Publishers): *Probleme der Curriculumentwicklung*, 2nd edition, Frankfurt am Main, 1972. Becker, H., Bonn, P. and Groddeck, N.: "Demokratisierung als Ideologie. Anmerkungen zur Curriculumentwicklung in Hessen". In: *Betrifft: Erziehung*, 1972, number 8, pages 19-29. "Gravierende kulturpolitische Fehlentscheidung", interview with Moos, G. and Klafki, W., In: *Betrifft: Erziehung*, 1972, number 9, pages 25-28.

(2) Cf. Klafki, W.: "Handlungsforschung in Schulfeld". In: *Zeitschrift für Pädagogik*, 1973, pages 487-516, especially pages 492 ff.

(3) Cf. the articles in *Beiträge zur Bildungstechnologie* (the journal published by the "Bildungstechnologisches Zentrum") in particular No. 3/1972 and No. 2/1973. Instead of the series "Project I: Untersuchung und Entwicklung einer lehrer-bezogenen Strategie für Curriculuminnovation und emanzipatorischen Medieneinsatz", hectographic copies of which have also been published by the Curriculum/Social Science Working Group, I refer to the synopsis of the project to be published shortly: Heinze, T., Müller, E., Stickelmann B. and Zinnecker T.: *Handlungsforschung im pädagogischen Feld*, Munich, 1975. Under the Coalition Agreement of the new Provincial Government of Hesse (SPD/FDP) signed in December 1974 the "Bildungstechnologisches Zentrum" is to be disbanded. It can therefore be taken for granted that the scheme of the above mentioned group will be either discontinued or else continued in a very abridged form.

necessarily belong together. On the one hand, decentralised curriculum development exists without any connection to research, especially action research⁽⁴⁾, and on the other, action research is not, as you know, necessarily related to curriculum development.⁽⁵⁾ However, I think it will become clear in the course of my remarks why action research must be combined with curriculum development as we understand its aims and tasks.

In what sense will I be using the terms "curriculum" and "curriculum development"? I use them in a wide sense.

The terms "curriculum" and "curriculum development" refer then to decisions and reasons for decisions on three levels, i.e.:

- Decisions and reasons for decisions as to
- general aims and special learning objectives
- subjects and contents of classroom instruction oriented to these aims and objectives
- the organisational forms, methods and media of classroom instruction oriented to the aims and objectives and appropriate to the subjects and contents.

The term "decentralised" implies that this kind of curriculum development takes place not only in co-operation with teachers but also in direct connection with their teaching practice, and if possible, with the co-operation of pupils and parents. The next step is to explain the term "action research" as I use it here; in this connection, I always mean: educational action research. This term can be temporarily defined by three general characteristics; some further characteristics will be mentioned later:

- a) In its interest in knowledge and thus in its questions, action research is related from the very outset to teaching practice, and this practice is understood as a specific form of social practice. Action research attempts to solve practical teaching problems. Thus it does not consider itself purely theoretical and neutral research.
- b) Action research is carried out in direct combination with the practical problem-solving attempts which it makes. As research it intervenes directly into practice, not waiting until the research process is completed; therefore, it must remain open to retroactive effects from this practice, which it itself helps to influence,

on the questions and research methods — during the research process itself, not waiting until the final evaluative stage with respect to future research.

- c) Action research must try consciously and purposely to eliminate the division between researchers and those involved in school practice. Its object is to put both groups in the innovation and research process into a relationship of direct co-operation and mutual influences.⁽⁶⁾

Basic problems of decentralised curriculum development in the form of action research

After this introduction, which served mainly to clarify terms, I now come in the main section of my lecture to the discussion of some basic problems of decentralised curriculum development in the form of action research. Some of these problems have already been briefly mentioned in the introduction.

As far as I know, no historical presentation of the development of the concept of decentralised curriculum development combined with action research exists, and I am not in a position to offer such a history here. One of the precursors of this new type of research is, as I understand it, Dewey's school reform work before and at the turn of the century.⁽⁷⁾ A further important step in America was taken by Hilda Taba's curriculum work and her beginnings of a theory of curriculum development.⁽⁸⁾ But of course, these are only very sporadic remarks.

In Europe, the first beginnings of a decentralised curriculum development, only parts of which, however, were and are accompanied by action research, have been observed primarily since the late 1960's probably most clearly in England and the Federal Republic, but also in Sweden and Switzerland.⁽⁹⁾

At least two basic motives are recognisable which have led to this development: One motive refers

(6) See the article mentioned in footnote (2).

(7) Cf. Bohnsack, F.: *Erziehung zur Demokratie. John Deweys Pädagogik und ihre Bedeutung für die Reform unserer Schule*. Unpublished thesis presented for the lecturer's qualifying examination, Marburg, 1974. A more extensive version is due to be published in book form by Otto Maier-Verlag, Ravensburg, 1976.

(8) Taba, H.: *Curriculum Development. Theory and Practice*, 3rd edition, New York, Chicago, San Francisco, Atlanta 1970.

(9) Cf. Gerbaulet S. et al.: *Schulnahe Curriculumentwicklung*, Stuttgart 1972. — Aregger, K.: *Interaktion im lehrerzentrierten Curriculumprozeß*, Weinheim/Basel 1973.

(4) Cf. for example Burst, R. et al.: *Weinheimer Gesamt-schul-Curricula*, Heidelberg, 1971.

(5) Cf. for example the majority of contributions to *Aktionsforschung*, published by Haag, F., Krüger, H., Schwärzel W. and Wildt, J., Munich, 1972.

to the failure of centrally developed, teacher-proof large curricula in the US. This failure was so alarming to curriculum theory and practice in England and the Federal Republic because many curriculum theories and some practical curriculum projects in these countries had tried to orient themselves to a large extent to American examples. The second motive is, I believe, more significant; it is a definitely political motive.

Everyone demanding consistent democratisation of society as a whole and the school system in particular, i.e., self-determination and co-determination, self-responsibility and co-responsibility, must also relate this principle to curriculum development and curriculum research. If curricula for the education of young people in a democratic society are supposed to be consistently oriented to the principle of democracy, that is if they are supposed to help young people acquire the ability to apply self-determination and co-determination, and if teachers are supposed to be in a position to teach in the sense of such curricula, then it would be unthinkable for these teachers not to be involved in the development of these curricula and the decisions put into them. Teachers who are not given the opportunity to help decide on aims, objectives, contents, organisational forms, methods and media of their classroom instruction cannot train young people to apply self-determination and co-determination!

This means nothing less than that curriculum development cannot be carried out according to a technocratic model "from top to bottom", not according to the model called engineering or research-development-dissemination model. It is the model which has been practiced since the 1950's and particularly since the early 1960's in the US and occasionally in Sweden. (10) Most of the forms in which curriculum development or guideline development have been carried out up to now in the Länder of the Federal Republic cannot be termed realisations of this research-development-dissemination model, because they usually lack the scientific aspect, that is the research aspect. But with regard to the characteristic "from top to bottom", they are still comparable to the RDD model. Perhaps we could speak of a model abridged by the research element, that is of a development-dissemination model. Here I cannot go into a description of the process of criticism of the engineering model for curriculum development and the counter-development of a non-technocratic model, a conception of curriculum development closely related to practice. I can only indicate a few names and institutions

which have played a part in it. In addition, I must emphasize that the stress on the political aspect of this countermovement varies widely from one author or institution to the next.

For America, we might mention beginnings in the sense of the so-called problem-solving model practiced, for example, at the Humanistic Learning Center in Albany, New York. In addition, there are in America — and especially in the United Kingdom — at least some projects on decentralised, open curriculum development within the Teacher's Centers Movement. (11) We could also mention the English Nuffield Junior Science project (12) and the approaches of Lawrence Stenhouse (13), which include some aspects of decentralised curriculum development.

In Sweden, the work within some of the so-called "pedagogical development blocks" formed in Malmö, since 1964, for example, is also being done along these lines. (14)

Perhaps it is only due to a lack of information that I have the impression that theoretical interest in problems of decentralised curriculum development and open curriculum appears farthest advanced in some groups and publications in the Federal Republic of Germany, partly because we have already been able to use international experience. First of all I should like to mention the excellent thesis prepared by a project group which was stimulated by the Founders' Association of German Science in 1972 under the title of "Schulnahe Curriculumentwicklung" (School-based Curriculum Development). (15) This thesis contains a very detailed proposal on the formation of regional pedagogical centers which should stimulate, support, co-ordinate and evaluate curriculum-development projects carried out by teachers in co-operation with research groups. Many suggestions in this thesis have been included in a detailed experts' opinion by the Education Commission of the German Educational Council under the title of "Promotion of

(11) Gerbaulet, S. et al: op. cit. pages 176 ff.

(12) Cf. *Nuffield Junior Science: Teachers Guide*, 1. and 2. edition, London/Glasgow, 1968. — *Nuffield Junior Science: Mammals in the classroom*, London/Glasgow, 1967. — Jung, W.: "Das Nuffield Junior Science Project". In: *Die Grundschule*, 1968, No. 3, pages 45 ff. — Lucas, M.: "Das Nuffield Junior Science Programm". In: *Neue Wege der Schulbiologie in englischen Unterrichtspraktiken*, published by R. Piefffer, Berlin, 1970. — Klewitz, E. and Mitzkat, H.: "Nuffield Junior Science Project — Didaktische Prinzipien und Beispiele". In: *Die Grundschule*, 1973, No. 3.

(13) Stenhouse, L.: "Curriculumentwicklung als Experiment". In: *Zeitschrift für Pädagogik*, 1973, S. 447 ff.

(14) Gerbaulet, S. et al: op. cit., pages 71 ff.

(15) Cf. footnote (8); Gerbaulet, S. et al.

(10) Cf. Gerbaulet, S. et al: *Schulnahe Curriculumentwicklung*, Stuttgart, 1972, in particular pages 50 ff. and 64 ff.

Practice-Related Curriculum Development", published in 1974.⁽¹⁶⁾ Since 1972, the "Curriculum Working Group" founded by scholarship receivers of the Volkswagen Foundation, which publishes a magazine, "Subject: Curriculum", devoted particularly to the problem of open curricula and showing much in common with this new type of decentralised curriculum development.⁽¹⁷⁾

In practice, decentralised curriculum development work in the Federal Republic has been mainly done in several Gesamtschulen, but only to a very limited extent with the support of research groups.⁽¹⁸⁾ The laboratory school and the Oberstufenkolleg at Bielefeld University, for which Hartmut von Hentig gave the decisive impulse for the set-up and development, are two of the few thorough, long-range projects on decentralised curriculum development in the Federal Republic.⁽¹⁹⁾

The first practical models, which, although they do not fully correspond to the suggestions contained in the aforementioned thesis and the opinion of the Educational Council, are nevertheless at least closely related to these. They are two projects called "Pilot Projects on Regional Advanced Education for Teachers" which have been running in Hesse since 1972. These projects concentrate mainly on formulating and testing the new Hessian Guidelines for Grundschulen and Secondary Level I. Here teacher groups cooperating on a voluntary basis are supported by small groups of scientists. The project group performing this experiment for the Ministry of Culture publishes a very interesting, practice-related magazine called "Curriculum Konkret" ("Concrete Curriculum").⁽²⁰⁾

Finally, I should like to name the Grundschule-project at Marburg University, already mentioned earlier, which I direct and which we regard expressly as a project on decentralised curriculum work in the form of action research, moreover the project at the Bildungstechnologisches Zentrum, Wiesbaden, also mentioned above.

Systematic aspects

After this brief outline dealing with the beginning

of decentralised curriculum development and the related action research, I shall now turn to the systematic aspects of the question.

All theoretical and practical beginnings which I have already mentioned include more or less criticism of the model of centralised and closed curricula which have been or will be developed according to the engineering model. From this criticism of the research-development-dissemination model, which is in contradiction with the democratisation principle, we can draw two positive consequences:

- a) Those active in school practice, particularly teachers, must be able to participate actively and decisively in the process of curriculum development. Increasingly, pupils and parents should also be included in this development; however, one must not ignore the difficulties presented by just this second demand.

We cannot, however, take for granted that teachers, pupils and parents possess from the very outset the competence, i.e., the perception, knowledge and abilities, necessary to take critical and productive part in curriculum development work. Anyone who proceeds from this false assumption will necessarily produce disappointment on the part of all concerned. The task is rather as follows: the curriculum development process must be considered at the same time as a *learning process* for all those involved. This applies not only to pupils and parents, but also to teachers and — as will be discussed later — to curriculum researchers. Here we are mainly interested in the situation of the teachers. Up to now, they have generally been able to acquire the qualifications for co-operation in curriculum development only to a very limited extent during their studies. But even if teacher training can confer more favorable starting positions in this regard in the future, the main requirement will still apply that the process of work on the development of curricula or curriculum parts (teaching units or sequences) must be organised as a learning process for all concerned.

- b) A second, positive consequence is the following: decentralised curriculum work must aim for open and variable curricula and curriculum parts⁽²¹⁾. It would be a contradiction in itself

(16) Deutscher Bildungsrat, Empfehlungen der Bildungskommission: *Zur Förderung schulnaher Curriculumentwicklung*, Stuttgart, 1974.

(17) *Thema Curriculum*, published by the Arbeitskreis Curriculum, since 1972, Verlag Lothar Rotsch, Bebenhausen.

(18) Cf. footnote (3).

(19) Hentig, H. von: *Das Bielefelder Oberstufenkolleg*, Stuttgart, 1971. By the same author: *Die Bielefelder Laborschule*, Stuttgart, 1971.

(20) *Curriculum Konkret. Dokumentation eines Modellversuchs*, published by the Projektgruppe Regionale Lehrerfortbildung, since 1973.

(21) Brügemann, H.: "Offene Curricula". In: *Zeitschrift für Pädagogik*, 1972, pages 93 ff. — By the same author: "Lernziele im offenen Curriculum". In: *Thema Curriculum*, 1972, No. 2, pages 16 ff. — Moser, H.: *Handlungsorientierte Curriculumforschung*, Weinheim, 1974. — Deutscher Bildungsrat: *Empfehlungen der Bildungskommission: Zur Förderung schulnaher Curriculumentwicklung*, Stuttgart, 1974, in particular pages 120 ff.

for a teacher group to want to produce, in co-operation with a group of researchers of curriculum packages as a result of their decentralised development work, i.e., products to be offered to other teachers and teacher groups simply for uncritical acceptance and application, like a recipe. On the other hand, we must emphasise that the "products" of decentralised curriculum development can only be examples, variable models for other teachers. They should, as examples, illustrate how certain objectives and subjects can be realised in the classroom under certain conditions given by the school, the pupils, the teachers and the concrete situation involved. They should stimulate other teachers not to schematic copying, but to *analogous* realisation. At the end of my remarks, I will return to this aspect.

Developing instruments and results of research

First of all, however, let us ask ourselves about consequences to be drawn from these remarks for research within the framework of curriculum development work.

Research which puts itself in the service of this kind of curriculum development is also decisively influenced by the democratisation principle just emphasised. It cannot regard itself as a neutral, exclusively descriptive or analytic discipline. In fact, it must be given the same aims as curriculum development. Its questions must be directed toward finding conditions, opportunities and difficulties in curriculum development as a democratic decision-making process. Research must be designed in such a way that it informs all those affected of the contents, pre-requisites and consequences of their decisions. Put more precisely, this means that research should, more than anything else, be an aid to *self-enlightenment* of those involved.

Action research in this sense is thus a special kind of "formative evaluation". Its results should be included as quickly as possible in the curriculum-development process in progress. I will give you an example. Let us assume that a group of researchers and teachers arrives in the course of a rather general discussion on the objectives of democratic teaching at the conviction that pupils should, among other things, obtain insight into the social interaction process in which they are involved in order to be able to act more consciously in such interactions. Now the group has the idea that one important aspect of interactions in school is the relationship of the pupils to the teacher. How do the pupils regard their teacher? As an authoritarian official? As a source of information? As too strict or too lenient? Are they afraid of him? What do

they expect of him? Which of his behavior patterns do they reject? When do they react positively to his suggestions? Do they see any contradictions between the aims and principles which he explicitly expounds and his actual behavior? And so on and so forth. For setting up a curriculum unit on this subject, it is centrally important to possess a scientific instrument for finding out the pupils' answers to the above questions. For example, a pupil questionnaire, an interview guideline, a category scheme for the evaluation of tape recordings of class discussions on these questions⁽²²⁾, perhaps even a projective instrument similar to a thematic apperception test. The results of such an inquiry, however, must be evaluated as soon as possible and put at the disposal of teachers and pupils as material for further analysis, discussion and changes in their interactions in school.

This small example already shows how difficult the task is of developing instruments and results of research which can contribute to the self-enlightenment of those involved. I will return to this problem later.

Developing open and variable curriculum parts

First of all, however, it is important to eliminate a possible misunderstanding with regard to the range and efficiency of decentralised curriculum development.

Some publications on decentralised curriculum development give the impression that decentralised work *alone* can solve the curriculum problem; decentralised curriculum development is then represented as the only alternative to all forms of centralised curriculum work. I consider this a serious error of grave consequence.

Decentralised curriculum development in the form of action research cannot all alone take over the task of developing general curricula for entire school systems, school types — all Gesamtschulen or all Gymnasien of one Land, for instance — or for entire school levels — the so-called orientation level for pupils aged ten to twelve, for instance — or curricula for entire fields of classroom instruction, such as science, or for single courses, such as history, should these encompass several years. The specific task of decentralised curriculum development in combination with action research consists, in my opinion, of developing open and variable curriculum parts, namely teaching units or sequences and strategies for such curriculum work.

(22) Cf. Haller, I. and Wolf, H.: "Selbstreflexion der Lerngruppe auf ihr eigenes Verhalten als didaktische Kategorie der politischen Bildung". In: Facklner, K.: *Handbuch des politischen Unterrichts*, Frankfurt am Main, 1972, pages 239-289.

Such a development of curriculum parts usually requires an existing general curriculum framework or even better, at least two such frameworks: on the one hand certain data or ideas on the larger context of the field or course into which the teaching unit or sequence under discussion belongs, and on the other data or ideas about the system of instruction fields, courses, etc., which is to be binding for a certain school system, school type, age level or grade. Unless we are to answer this question in completely amateur fashion or with uncritical acceptance of traditions, we are expecting too much of a number of small groups of teachers and a few action researchers practising decentralised curriculum development. Here, it is necessary to have information on the status of general curriculum theory and didactics, of the field or course under discussion, on the related scientific disciplines etc.; it would be impossible for each group working on decentralised curriculum development to work these out alone. For this reason, decentralised curriculum development must relate to supra-regional institutions responsible for an entire district, a land or a whole state, institutions which would then develop more comprehensive extensive general, field and course curricula (Rahmen-Richtlinien) ⁽²³⁾. However, our earlier remarks lead to an important consequence: it is clear that even such comprehensive curricula cannot be teacher-proof curricula or strict, absolutely binding lesson plans (Lehrpläne). Then decentralised curriculum work could only put the final concrete touch on pre-set instructions or fill out an unalterable frame. If, on the other hand, decentralised curriculum development in the above sense is to be a creative activity and a kind of co-determination of teachers and — if possible — pupils and parents on aims objectives, contents, organisational forms, methods and media of classroom instruction, then the more general data must have the form of open curriculum guidelines or open curriculum systems. "Open" means here that these general guidelines must be criticisable and alterable; they must be regarded, so to speak, as normative hypotheses which can and should be put into concrete form and tested in the classroom and designed, checked, changed, supplemented for varying situations. A suitable form for such an interpretation, in my opinion, is "schoolbased curriculum development", combined with action research.

This concept is, therefore, in opposition to two contrary models on the relationship of general

curriculum decisions on the one hand and the individual elements of a curriculum on the other.

The first contrary model I call the "inductive model": I mean the idea that a comprehensive curriculum can be created by the addition of many individual, decentrally developed curriculum elements.

The second model I call the model of "deductive curriculum development". It is based on the idea that some central curriculum authorities — parliaments, central commissions, scientific curriculum institutes or whatever — can set up general learning aims and objectives and curriculum principles clearly and bindingly, but that the further process of curriculum development consists of deduction from that general aims objectives and principles until one finally arrives at concrete teaching units and sequences.

As opposed to this, the relationship between central and decentralised curriculum development must be defined differently, as a relationship of open interaction, as a process of constant, reciprocal interpretation, correction and advanced development. Of course, I am not describing a given situation, but am expressing a demand which is, in my opinion, justified.

Why must decentralised curriculum development be considered a learning process?

It has already been emphasised above, that decentralised curriculum development in the form of action research can be successful only if it is regarded as a learning process for all affected groups and organised accordingly. I will now return to this point. Why must decentralised curriculum development be considered a learning process? I will discuss this question using the example of the teachers and members of action research groups. Both groups, teachers and action researchers, usually enter the working process of decentralised curriculum development with varying qualifications and pre-requisites. They set themselves the task of solving practical problems together, developing curriculum parts and realising them in the classroom and of critically processing the experiences thus gained. But, it is not all a matter of course that both groups regard and interpret the tasks and problems in the same way from the very beginning. Here are some examples ⁽²⁴⁾:

(23) See also: Deutscher Bildungsrat, Empfehlungen der Bildungskommission: *Zur Förderung schulnaher Curriculumentwicklung*, Stuttgart, 1974, in particular pages 25 ff.

(24) Cf. Klafki, W.: "Handlungsforschung im Schulfeld". In: *Zeitschrift für Pädagogik*, 1973, in particular pages 502 ff. — By the same author: "Probleme des Zusammenwirkens von Wissenschaftlern und Praktikern". In: *Zeitschrift für Pädagogik*, 1973, pp. 91 ff.

Teachers do not generally have sufficient training in scientific research methods. They are forced to teach every day and to make constant decisions on short notice. They must act within a complex field of relationships, demands, tasks and interactions and assert themselves successfully every day: in their relationships to the children, to their colleagues, to parents, school administration, etc. They expect co-operation with a research group to help them rapidly with their numerous difficulties in everyday school life, and they hope for relatively quick, noticeable improvement in the quality of their teaching.

The members of a research group, however, do not always have teacher training or practical school experience. But, even when they do have this training and experience, they are not under the daily pressure to act to which teachers are exposed. This alters their perspective of school reality. Their questions are usually of a longer term, but also more selective. They see only certain aspects of the complex school reality, but these usually in finer detail and with more theoretical reflection than those involved in school practice. They expect the form of their co-operation with teachers and the results of this work to correspond as far as possible to scientific standards — and this means first of all, to theoretical standards. These varying perspectives — although both sides usually have the best of intentions and although they want to work on a task together — often lead to group-dynamic tensions, disappointments and conflicts. I will name two dimensions in which such tensions very often become apparent:

- a) in the different languages spoken by the two groups;
- b) in the different expectations they have of one another.

In both dimensions, therefore, the joint curriculum-development process must also be a learning process:

- a) each group must learn to understand the language of the other and try to find a language which both sides can understand;
- b) each group must learn to understand what the other expects from this co-operation and why, and they must learn to harmonize their varying expectations in order to solve their common tasks.

In German action research theory the discussion has often arisen that between the various partial groups co-operating on an action research project there must be no strict distribution of roles and

functions or at least that the difference in roles and functions apparent at the beginning must gradually be eliminated. I believe that this is an unrealistic demand in most cases. The thought behind it, however, is correct. I think it can be expressed in three postulates:

- a) All members of a group of teachers and action researchers must learn to *understand* the various roles and functions to be filled in such a project.
- b) The borders between the various roles and functions should be kept as flexible as possible. For example, it would appear desirable for the "researchers" in a decentralised curriculum-development project to teach occasionally and for teachers, or even pupils and parents, to help work on the development of research instruments, perhaps a questionnaire and then to use this instrument themselves, etc.
- c) Even when the various groups of persons assume differing roles and functions in such a project for practical reasons, these roles and functions must be recognised as *equally important*.

Evaluation or use of scientific methods

These reflections, now, have consequences on the development of methods and instruments of research.

The problem of evaluation or the use of scientific methods must be re-formulated within the framework of decentralised curriculum development and action research in relation to classic empirical research, not only on the level of general methodological principles, but, also on the level of development or application of particular methods. First of all, a word as to the methodological problem level: At least two criteria which have up to the present been regarded as binding for classic empirical research cannot be applied to action research:

- The clear division between the researcher and his research instruments on the one hand and the research object on the other (teachers, pupils, interaction between teachers and pupils in the classroom, etc.) can no longer be accepted if action research is supposed above all to serve the self-enlightenment of all those involved. The researchers must not turn the investigated persons and their relationships into simple objects of research. They must not, for instance, keep the objectives of the research or the research methods a secret from them.

On the contrary, the questions of research and — as far as possible — the research instruments should be discussed as often as possible with the teachers or parents or pupils *before* they are applied and the results should be put at their disposal as soon as possible. I know myself, partly from the Grundschole-project which I direct, how difficult it is to realise this principle and how far removed we remain from it in practice.

- The second criteria concerns the requirement that the researchers must be able to repeat an investigation under the same conditions and that in a complex field such as classroom instruction only a few variables must be selected and studied, whereas the others must be kept constant over a long period of time. This principle holds, if at all, to a very limited extent in school-based curriculum development and action research in our sense of the word. In such projects, it is *change* in complex structure which is to be planned, tested and analysed. Suppose that a teacher-researcher group wants to develop a teaching unit or a sequence of such units for elementary science instruction in the primary school⁽²⁵⁾. The unit is to be related to the problems and to the interests of the children and their experiences with their environment, and the children are to become acquainted with elementary scientific methods in order to answer their questions: observation of animal behaviour, little systematic animal experiments, etc. In this instruction, however, they are also supposed to reflect on their own relationship to animals. With such a complex objective, no individual variables — objective variables or thematic variables or methodic variables — can be isolated and changed whereas others are set, either in planning or in execution. And in the second run of an analogous planning or teaching process, in turn, we cannot rigidly attempt to keep the existing conditions and reject, for instance, creative suggestions of the children which might bring about a radical change in the teaching unit only in order to permit repetition of classroom observations under the same conditions as in the first experiment.

If we accept these remarks as valid, however, we are confronted with a double question which I can only pose here, but not answer satisfactorily:

- a) How, then, are scientific statements possible, or: must the term "validity of a scientific statement" be redefined?
- b) How can knowledge "gained in one place" be generalised or transposed to other schools and classroom situations?

We now turn to the level of individual scientific processes or instruments for action research as part of decentralised curriculum development. Here as well, the ground is hardly broken. In this lecture, it would be impossible to discuss individual methods and detailed problems. Here we can only name a few basic aspects and difficulties arising in the development of appropriate research methods and instruments as part of action research; some of these aspects have already been briefly mentioned. I will name four aspects already presented in my paper:

- a) The main questions, structure and results of the research processes should be easily understandable to those involved. Only under this condition they can contribute to the common learning process of all participants.
- b) A second difficulty in the development of useful research instruments for action research arises from the guiding objectives for development of curriculum parts. The objectives of teaching units or sequences oriented to the gradual development of self-determination and co-determination of the children cannot as a rule be operationalised as individual, isolated types of behaviour in the sense of behavioural psychology. Along the lines proposed by R. Mager⁽²⁶⁾ the following objectives can be mentioned:
 - the cooperation ability of the children,
 - the ability to help themselves in learning and working processes,
 - the ability to introduce their own ideas and interests into discussion, etc.

These are abilities which can be proved learned or not learned only in the context of certain situations and in classroom and extra-classroom processes; therefore, they can be scientifically observed only in the context of such situations and processes. But, who has the suitable research instruments?

(25) I refer here to a teaching unit developed by Mrs. Ch. Bödecker in co-operation with teachers as part of the Marburg Primary School Project.

(26) Cf. Mager, R.: *Lernziele und programmierter Unterricht*, 14th edition, Weinheim, 1970. — By the same author: *Zielanalyse*, Weinheim, 1973.

Bales, Flanders, Bellack, and others⁽²⁷⁾ show the beginnings of scientific analysis of classroom processes, but, they are still too formal and must be developed further — with respect to aims and contents, as indicated above.

- c) In a type of education and instruction oriented to the aim and objective of self-determination and co-determination, the interest of research must be directed not only towards such complex aims and objectives as those named above, but also to the *conditions* and *processes* under which such "learning results" are achieved or prevented. How must classroom situations be structured in order for children to develop communication abilities? What kind of learning processes motivate children to develop the ability to criticise social prejudice?
- d) A fourth requirement of appropriate methods for action research has already been mentioned earlier, but, we must turn to it again in this context. If research in the process of school-based curriculum development is to offer teachers and pupils the help to attain such aims and objectives as mentioned above, then it must develop instruments which permit at least partly quick feed-back of results — or partial and intermediate results — to the teachers and pupils. Only under this condition can they, supported by action researchers, draw consequences from the reports for further teaching and learning processes during the process itself. Here again, we must emphasise how difficult it is to meet this demand, if these intermediate research results are to be in the least reliable.

(27) Bales R.: "Die Interaktionsanalyse". In: *Beobachtung und Experiment in der Sozialforschung*, published by René König, 3rd edition, Cologne, 1969, pages 148 ff. — Flanders, N. A.: *Helping teachers change their behavior*, Michigan, 1963. — By the same author: *Teacher influence, pupil attitudes and achievement. Co-operative research monograph*, No. Ber 60-6, Minneapolis, 1965. — By the same author: *In-achievement. Co-operative Research Monograph*, No. Ber 60-6, Minneapolis, 1965. — By the same author: *Interaction analysis in the classroom*, Ann Arbor, 1966. — Frech, H. W.: *Kontrollierte Beobachtung verbaler Verhaltensweisen im Unterricht*, Neue Sammlung, 1971, pages 87-108. — Bellack, A. et al: *Die Sprache im Klassenzimmer*, Düsseldorf 1974. — Cf. also Koskenniemi, M.: *Elemente der Unterrichtstheorie*, Munich, 1971.

Open and variable curriculum products

I return to a question already mentioned: What nature can or must have the "result" of decentralised curriculum development created with the aid of action research? The fact that the results cannot be teacher-proof curricula or teacher-proof units to be re-distributed as "pre-fabs" to other schools, have already been made clear by teachers and pupils. On the other hand, however, there are neither sufficient theoretical nor practical reasons for assuming that decentralised curriculum work in the form of action research cannot create any "products" which could be transposed in any sense to other schools, teachers, classes, and learning situations.

The purpose of open, variable curriculum products (teaching units or sequences) arising from decentralised curriculum development by certain groups of teachers and action researchers must be to stimulate other teachers, groups of teachers and teacher-pupil groups to an *analogous* form of teaching and learning and to their own curriculum-development work⁽²⁸⁾. Such "products" of individual decentralised teacher-researcher groups, that is teaching units or sequences or strategies of co-operative lesson planning, might be adapted by other teachers or groups of teachers under the following conditions:

- a) Checking and criticism of the products with respect to their aims objectives and prerequisites must be possible. This means that not only formulations of learning objectives, structuring of subjects, etc., must be represented in such a curriculum product, but primarily the *reasons* found by the creating teacher-researcher group for their decisions as to learning objectives, subjects, contents, methods and media, and the *conditions* which they took as a basis: for instance, the starting conditions for learning and the socio-cultural situation of the children, the teachers' own estimation of their abilities, etc.
- b) The products should be illustrated by examples, that is they should contain concrete examples of discussions and learning objectives within the teacher-researcher group, for instance, or examples of pupils' working materials, extracts from teaching records, pupils' work, etc.

(28) Cf. also: Giel, K. and Hiller, G.: "Verfahren zur Konstruktion von Unterrichtsmodellen als Teilaspekt einer konkreten Curriculumreform". In: *Zeitschrift für Pädagogik*, 1970, pages 739-754. — Hiller, G.: *Konstruktive Didaktik*, Düsseldorf, 1973.

- c) The products should contain at least two to three practiced variants of the teaching unit or sequence under discussion. As examples, these variations should show how the same teaching objective complex — for instance, "childrens' experiences with animals and reflections on their relationship to animals" — might be realised in various ways in the classroom according to differing conditions in the school, the class and the teacher or group of teachers.
- d) The products should also open and understandably discuss any difficulties and errors recognized during evaluation.

Role of the teacher in school and curriculum reform

I have attempted to develop one basic thought all through this lecture: the central significance of the teacher in the process of school and curriculum reform. In conclusion, one special aspect of the situation of the teacher in the process of decentralised curriculum development in the form of action research must yet be emphasised. Instruction corresponding to the aims implied above makes great demands on the teaching abilities of the teacher. Now, this fact has one consequence which presents new problems.

First of all, teachers in the process of decentralised curriculum work must develop a high degree of knowledge and reflective ability with respect to the various partial aspects of instruction; second,

their ability to plan lessons is increased. This, however, gives rise to one difficulty: even outside of decentralised curriculum development, a considerable discrepancy can frequently be found, between the demands teachers make on their own instruction — that is, their ideas of how they would like to teach — and their actual teaching practice. But, this does not at all mean that the actual teaching qualification of the teacher automatically increases. Teachers often find the tension between their own demands and wishes on their instruction on the one hand, and their teaching reality and actual teaching qualifications on the other, depressing. Therefore, the demand is frequently heard that curriculum development should be combined with new forms of practical advanced teacher training, with in-service training closely related to teaching practice. Without doubt, this is a correct postulate. But, as long as the present planning of open and dynamic teaching units and sequences remains unsolved along the lines outlined above, it would be very difficult to attack the problem of in-service training. In the Marburg Grundschule-project, for example, this double demand is a problem which can hardly be overcome. For advanced teacher training closely related to practice, there are hardly any proven methods or instruments, with which a systematic increase in teacher qualification for actual teaching might be obtained and which could be practiced under normal school conditions. Further action-research projects should try to develop such methods and instruments from the very outset parallel to the methods of formative curriculum evaluation within the framework of decentralised curriculum development.

School administration and evaluation of pilot projects

by E. VOIGT,
Director of the Centre of Education, Berlin.

PART I

The subject relates two very general concepts, to each other, namely school administration and the evaluation of school reform pilot projects. Both concepts cover a number of quite different situations.

School administration, as L. Legrand convincingly shows in his 1973 report, covers different functions and fields in the various European countries. This means, in my experience in the Federal Republic

of Germany, that federal government has few powers, perhaps being of consequence only as a source of finance, while in each Land there is parliamentary responsibility and the administration system is organised to meet its requirements.

The evaluation of pilot projects in different Länder varies with their attitudes towards the scientific comparison of pilot schools with traditional schools. Länder with Christian Democratic Union (CDU) governments demand such comparison and base further school reforms on scientifically controlled

experience. Länder governed by the Social Democratic Party (SPD) adopt the view of the majority of scientists that such a comparison cannot be carried out methodically. They also think this an unreasonable demand on the schools, as it leads to competition which is familiar enough in commerce but would be unprecedented in school history. They have no difficulty in deciding against comparison, since the changes sought are part of the political programme anyway.

The administration may have various functions in connection with the evaluation of pilot projects:

- it may be interested in obtaining answers to questions or analysing complex situations,
- it may be responsible for applying and administering public funds for such projects,
- it may act as the official body authorising questionnaires for pupils,
- it may have to exert its authority over schools which have reservations with regard to scientific investigations or may have to resolve disputes between schools and education researchers,
- it may be in rivalry with, or even opposition, to research groups in assessing disturbing factors in school experiments and influencing teachers and the public.

Despite this multiplicity of functions, the administration is obliged by its structure to adopt a relatively clear-cut and reliable attitude. It has to speak with a single voice, and because of its organisation it is less affected by changes in staff than, for instance, project groups are. It must have a constant presence: it cannot, retire or change its attitude overnight. All statements for which it is answerable to parliament acquire the force of medium-term commitments, at least for that parliament's lifetime. The citizens are entitled to such commitments, being in any case unsettled by the excessively swift changes in education. But such commitments are not welcomed by education researchers, who want their often important interim results to be taken into consideration immediately.

Evaluation of pilot projects has already become very heterogeneous in practice:

- Frequently data are collected and evaluated in the schools, on the basis of specific hypotheses, by teachers with a scientific training or teachers assisted by scientists. They are under the authority of the school and are also subject to the constraints of the school's everyday life. This is particularly effective

in the case of long-term, continuous studies. The studies by Teschner and Fischler in Berlin are examples of this type of evaluation. It would be wrong to expect only concordant results from such groups. It is to them that we owe our knowledge about the important selective function of options in compulsory education, the dangers of thoughtless action based on differences in achievement and the prospects for heterogeneous study groups in the natural sciences, although their findings do not correspond to their original hypotheses or those of their administrative authorities.

- Studies by outside institutions or groups are more independent, but in this case continuous contact with the school is more difficult. Usually such groups have the greatest chances of success in cross-section studies carried out at a specific time. But such studies are subject to the risks inherent in assembling data once and for all in their final form. In spite of its independence, this type of study does not present the administration with any problems, even if the results are often unwelcome. The clear structure of the process (construction of hypotheses and collection, evaluation and interpretation of data) means that it can be used and checked by different individuals.

Disputes usually occur over the approval of questionnaires to schoolchildren; whenever these are used in compulsory state schools they must be justified in educational and legal terms before the public and parliament.

- The mixed groups of scientists and teachers in more recent projects are a new phenomenon which the administration has not yet completely explored. In the new "action research" pattern teachers have equal rights; knowledge acquired concerning a field of action is translated directly into concepts to bring practical influence to bear on that field. I have no experience yet of the effectiveness of such projects as far as transferable results are concerned. Personal reports are usually rather too absorbed with detail and too euphoric; they remind me of the reports on school outings I used to submit to my authorities when I was a class teacher.

Because of its function and its very nature the administration is bound to regard such forms with suspicion. It is true that this pattern is attractive at first sight: science comes down off its pedestal and exposes itself to the test of practice. In fact,

however, it acquires new powers: it brings its intentions to bear in interactions and feedback without having to defend them publicly, with the risks that entails (e.g. loss of an election). Often there are general arrangements with the administration which, precisely because they are so general, are difficult to relate to specific situations.

The remarkably limited nature of such projects, being confined to a handful of participants and schools, makes action research very unobtrusive politically, it is true; but is it in fact realistic to ignore the administration, the aims of the parliamentary majority and scientific discussion? To apply findings directly in the finders' own field of action, leaving all controlling bodies (the scientific public and government departments) out of account, does indeed give them the feeling that they are making progress, but also a sense of isolation in a world which they view as increasingly unsympathetic.

It is already clear that these mixed forms engender solidarity that is scientifically open to question. The scientists identify themselves with the "base" to such an extent that the detachment needed for the acquisition of knowledge is no longer guaranteed. The sense of dialectic between solidarity and knowledge is repressed. The scientist is happy to be accepted. Since his function in the school is in any case uncertain and unnatural, he finds his identity by accepting for himself the teacher's mission as something good by nature and distorted only by circumstances. "Teacher", "pupil" and "base" become almost sublime metaphors in the work of such project groups.

PART II

After this introduction I should like to describe the administration's expectations and experiences of the evaluation of pilot projects. First let me explain that I have become mildly sceptical in this respect although I, perhaps more than others, favour scientific research into schools and teaching as well as scientific thought on the part of the practising teacher. My scepticism concerns the particular case which you are discussing in this symposium: the marriage of school projects with scientific interest, which I myself have considered such a happy one in past years. I postulate seven theses:

1. School pilot projects have different values in the educational policies of the Länder. These values determine the prospects of success, and the results of evaluation, though this influence is not always realised or openly acknowledged.

2. In pilot projects the direct effects (those that can be communicated and are experienced by a large number of the people concerned) are more significant than the achievements and results of evaluation.

3. The education researcher is more likely to become involved in conflicts of loyalty in school projects than in other educational institutions.

4. Scientific research is bound to make for greater complexity. In school projects this will be a disturbing factor. The interests of the administration and the teachers demand that such complexity be minimised.

5. School reform, pilot projects and scientific evaluation have different time scales. The more closely science ties itself to with practice, the more it will disappoint all concerned, by reason of its virtually unchanged temporal requirements.

6. In practice the administration and teachers will always make selective use of the results of scientific evaluation. This is disappointing to research groups and gives rise to protests.

7. Model schools cannot be exploited indefinitely for research purposes.

Thesis 1: School pilot projects have different values in the educational policies of the Länder. These values determine the prospects of success and the results of evaluation, though this influence is not always realised or openly acknowledged.

Generally speaking, educational pilot projects are not genuine experiments.

— In Länder which intend reforms, even such projects have political significance as the outriders of a development which is desired in general terms. Scientific evaluation will forfeit its influence if with its modest instruments and staff it challenges the political premises. Its role will be to work on part problems and help to clear up many of the questions and constraints inherent in such reform processes.

— In Länder that seek to preserve the status quo because of their political majorities, pilot projects are often concessions to groups that are not satisfied with educational policy. First and foremost they are show-pieces. Such show-pieces also exist in Länder that seek educational change, their function being to placate conservative groups (in addition to 6-year elementary schools there have been

in Berlin since 1952 two classical grammar schools beginning in the 5th year; however, they have never been scientifically evaluated).

Progressive show-pieces raise special problems for the researcher, their situation being particularly difficult. More than in the case of other projects, the researcher will seek to support such isolated institutions and help to justify them.

— Honesty demands that we also mention a third group: pilot projects initiated in order to attract money to the school. Since, under the administrative agreement on pilot projects, the federal government provides funds for such projects in schools, these are often a way of raising the poor quality of school life. As our children benefit from the material improvements, I see nothing dishonourable in this.

The educational researcher whom fate assigns to such a project must try to give form and meaning to this vague urge for progress. In view of the slight political significance of such projects they may prove to be a wide field in which to gain knowledge of the subsidiary structures of school life.

Thesis 2: In pilot projects the direct effects (those that can be communicated and are experienced by a large number of the people concerned) are more significant than the achievements and results of evaluation.

The social effect of school pilot projects is direct rather than dependent on the results of scientific evaluation.

Waiting lists for enrolment in a school, the identification of parents and pupils with it, its "display value", the openness of its life to the mass media — all these are factors which generate their own dynamic. An experimental school has to stand the test of its immediate surroundings. As an institution with a difference it has to present its environment with a simplified image of itself and ensure that this self-imposed image corresponds with what visitors see. In view of their progressive aims, many experimental schools thus fall into a forced rigidity which makes it impossible to deal productively with unexpected hitches and new developments. The headmaster of many a model school finds himself in this situation. He is constantly having to project an image of his school and reacts sharply if research questions this image. Evaluation of school reform projects is research in a field full of tension and problems.

Thesis 3: The education researcher is more likely to become involved in conflicts of loyalty in school projects than in other educational institutions.

Usually evaluation of school reform projects covers individually identifiable groups. Seldom is the population so large that single schools can no longer be identified. If a single person or specific group is identifiable, problems of loyalty arise, for the observed behaviour of such a group is inevitably imperfect and inconsistent in some respects. I know of cases in which the school authorities and the immediate neighbourhood were easily able to identify the persons referred to in the research reports. In the case of a large anonymous population (e.g. teachers in special schools), statements based on empirical studies can generally be made which may be extremely effective politically without any suggestion that a particular person is to blame or has failed. The research results give rise to a discussion in which no individual is obliged to stand out.

Evaluation thus involves crucial considerations of possible disadvantages to identifiable participants. In my experience research groups have taken this very seriously and an administration anxious to support commitment and willingness to experiment in its school system will welcome such thoughtful consideration on the part of researchers.

In recent years it has become more usual to name deficiencies, while exonerating identifiable persons concerned by attributing the shortcomings to external factors. Studies thus end with demands for less teaching time, better equipment and greater independence for teachers. However, they also refer to factors about which nothing can be done, such as the parents' work or the fundamental clash between capital and labour. To such demands and statements, most of which are plausible, it can only be said that they are not the result of research. There is no proof that by meeting the demands the shortcomings could be prevented. Such demands are a moral appendage; they result from the need for loyalty to an identifiable group.

Thesis 4: Scientific research is bound to make for greater complexity. In school projects this will be a disturbing factor. The interests of the administration and the teacher demand that such complexity be minimised.

Like all research, evaluation of school reform projects makes for greater complexity. It shows that the school and its projects are different from what was originally supposed; they are more complex and more interdependent. Often the results are completely at variance with some naïve view. To take an example, in the sixties when in the Federal Republic of Germany the traditional school system was still being projected into a stable future and school reform was still just a subject for discussion,

administrations and teachers were both astonishingly open-minded towards the increasing complexity emerging from educational research. People accepted criticism of everything and even of themselves. It was a golden age for educational research although for lack of funds it was not possible to derive the full benefit from it.

This openness changed when reforms started; for the situation itself became more complex because of the co-existence of traditional schools, general reforms (e.g. guidance periods) and a large number of pilot projects. Further scientific findings which reveal new factors and again challenge ideas already put into practice are provoking increasing displeasure or being ignored. The situation is worsened by the very natural professional interest on the part of researchers to make a name for themselves by achieving new and revolutionary results.

The results of evaluation will retain their influence only if reform processes accept a number of decisions or traditions and does not ask too much of those concerned. Any rigid pursuit of research that disregards the need for such moderation in its demands may perhaps bear fruit in the year 2000; but I predict — though you may of course disagree — that it will have no effect in the years immediately ahead.

That is true of administrations that are confronted — as they often are — with factors over which they have no control. It is also true of teachers.

Discussions in recent years have resulted in changed standards and additional ones. While the "field of duty" is extending, the individual's capacity to achieve anything in his few working hours is still basically the same. The incompatibility between more complex norms and imperfect everyday life presents job satisfaction; indeed, it makes for susceptibility to ideas that offer hope of relief.

Thesis 5: School reform, pilot projects and scientific evaluation have different time scales. The more closely science ties itself to practice the more it will disappoint all concerned, by reason of its virtually unchanged temporal requirements.

This thesis does not say anything new. When research takes place remote from practice and the daily need for decisions, such differences do not become a public problem. The closer research comes to practice, the more it arouses hopes of a secure basis for decisions and the less fitted for this it comes to appear to practising teachers because of all the time it requires. Research groups

which often co-operate with practising teachers must find this reproach, whether explicit or implicit, discouraging.

The impression that the evaluation of school reform projects does not take time sufficiently into account is strengthened by the considerable amount of time which the project groups take to settle their internal problems. Such a group is a new social phenomenon of which we have very little experience. Its social equilibrium is very unstable: it tends to the formation of splinter groups and gives rise to many internal problems. There are times when it is largely inward-looking. Such time-consuming forms of decision-taking make a terrible impression on teachers working time-tables imposed by others. Indeed, the different time-patterns of scientists encourage the feeling in the administration and among teachers of a difference in the very quality of life.

Thesis 6: In practice the administration and teachers will almost make selective use of the results of scientific evaluation. This is disappointing to research groups and gives rise to protests.

Administrations and teachers are always being accused by researchers of selecting for practical implementation only those research results that suit them, and of never taking into consideration the findings as a whole, with all their "ifs" and "buts". This reproach is only partly justified. Nobody nowadays can suppress really serious objections, even if he wanted to.

I refer here to the justified criticism that a person who has to take a decision makes a selection, even if only, unconsciously. In my view this cannot be changed, and the strategy of evaluation must take account of the fact. The progress in education over the last centuries is not the result of implementing ideas and scientific findings in their entirety. In every case only a few factors had any historical effect. The premisses and the context, everything that was important to the educationalist and the research worker, were soon forgotten.

Thesis 7: Model schools cannot be exploited indefinitely for research purposes.

In recent years the administration has increasingly often found itself obliged to mediate between model schools and research groups, usually because of the unwillingness of schools to be subjected to research any longer. Last year in Berlin the Senator had to use his authority to ensure the implementation of another schools research project. Chris Argyris once made the pointed observation, "For-

merly the barons of the timber trade had trees felled without bothering very much about the future of the industry. Researchers use persons for experiments without bothering very much about the future supply of such people" (in "Gruppendynamik", 1972, Vol. 1).

Since moreover many students carry out empirical research for their state examinations and degrees and find educational pilot studies especially attractive for the purpose, it falls to the state to regulate the pressure of research on the schools.

The ideas I have put forward are the result of limited experience, and because of my job, are one-sided. I nevertheless draw the following conclusions from them:

— Evaluation is not just a problem of research methods. That became clear in 1970, when a large number of conflicts of interest and misunderstandings arose between schools, administrations and scientists.

— These disputes cannot be solved through systems of co-operation alone. There has been a clearing commission for such disputes in Berlin since 1972. That is a step forward. But even so there is a dogfight of expectations and interests because those concerned have no time to give thought to evaluative research as an instrument with all its opportunities and limitations. It is just not enough simply to get people with different interests around a table, however essential that may be.

— Evaluation also calls for realistic ideas on strategy. Special consideration must be given to:

- the Land's educational policy,
- the project's stability or instability,
- reducing complexity,
- the establishment of a realistic time-scale,
- the limit to the demands that may be made on schools.

From centralised to decentralised innovation – Problems of evaluation

by O. LINDAL,
College of Education, Stavanger.

The process of educational innovation — centralised as well as decentralised — may be regarded as a continuous transition through distinct stages, each of which involves dissimilar objectives, tasks and priorities. This holds for wide ranging educational reforms, as well as for specific projects.

These stages may be broadly characterised as:

- Planning
- Development
- Follow-up

Evaluation is seen as a vital element at each stage, if central, regional and local dissatisfaction with "ad hoc", random experimentation is to be avoided.

The tradition of centrally sponsored research

A strong tradition of political influence in educational planning is typical in those countries which possess a more or less monolithic school structure. In such countries, there has been a corresponding

tendency to rely on central political and administrative initiative for educational development. Where this is the case, it has fallen to central authorities to ensure evaluation of such schemes as have been put into effect. However, evaluation has not traditionally been considered an integral part of the process whereby funds are allocated for educational research. The bodies responsible for financing or sponsoring research projects have played a largely "passive" role. It has been taken for granted that once funds are allocated, it should be left to professional researchers to tackle the actual work involved in their own way, and to return reports and accounts on completion of any given project. Such studies have been largely carried out at universities under the leadership of established and experienced researchers according to relatively narrowly defined criteria.

Recent years have seen a distinct tendency towards direct and active involvement in the research process on the part of the sponsors themselves. Such involvement has included the circulation of infor-

mation acquired as well as distribution of materials developed in the course of projects.

Manifold educational reforms over the last two decades — organisational, structural and curricular — have prompted a political initiative from both inside and outside the educational system, leading to the creation of bodies specifically responsible for systematic evaluation of research projects. These trends were particularly apparent during the sixties. The resources allocated to these bodies were directed mostly towards policy orientated research, while the authorities themselves showed a growing interest in research findings.

Official interest has no doubt been prompted by a felt need for assistance when making decisions on complex educational matters.

Problems of centrally sponsored research

Centrally sponsored research, of the kind sketched above, has encountered a variety of problems, of which the following should be briefly mentioned as being relevant to attempts to organise research on a decentralised basis.

- The monitoring of projects leads, too easily, to interference in the conduct of the projects.
- Whenever research funds are allocated by agencies whose own research policy is clear cut, the conduct of the research itself tends to be dictated by those agencies rather than by the research staff themselves.
- It is frequently difficult to ensure adequate financing of research directed towards
 - critical evaluation
 - alternative models
 - evaluation of underlying basic assumptions.
- Central authorities have tended to give priority to short-term projects dealing with immediate practical problems.
- Funds are often allocated on a year to year basis, owing to the unwillingness of sponsors to commit themselves to supporting research projects until they are fully completed.
- It has, to some extent, been difficult to draw on the most experienced and best qualified research expertise for applied research.

Nevertheless authorities have in general attempted to create an environment in which the findings of educational research can influence educational policy as well as teaching. Authorities must be pre-

pared to take some risks if such influence is to be really effective. Countries with centralised educational structures have not yet had sufficient time to achieve wholly satisfactory solutions to the problems which have arisen during the course of this process, or to establish a sufficiently effective relationship between research and educational policy making. It is, in any event, certain that the effect of research findings on the decision making process in education must depend on the existence of an appropriate system for channelling relevant information.

The problems involved must be faced squarely, tackled seriously and, where necessary, unconventional solutions to them must be examined.

The experience of decentralisation in Norway

Recent experience in Norway may, perhaps, illustrate some of the problems. In Norway, it has become obvious that centrally directed attempts at research into educational questions are frequently doomed to frustration. There is now an increasing tendency to move away from centrally controlled reform and towards involving various types of regional and local institutions at different levels. Policies directed towards decentralisation are also characterised by a determination to employ more resources in the solution of educational problems, rather than towards supplanting central resources. Until recently, uniform educational structure and political determination have been regarded as the most important pre-conditions for raising the general level of educational opportunities for people from every section of society.

In the late sixties and early seventies some countries have begun to feel that the level of general education is sufficiently high (or is in the process of becoming so) for local or regional initiative to be relied on to develop further differentiated educational opportunities for groups and individuals in each area. In Norway for example, the state, for its part, has maintained financial support for the various types of school that have been established. In place of detailed instructions for, for example, standard curricula within a uniform system, central authorities are increasing favouring indicative norms. At the same time it is admitted that the success of this policy must, to a large extent, depend on local and regional initiatives and the active participation of teachers, parents and pupils in the organisation and provision of educational choices. Such initiatives should also be able to contribute to the establishment of special courses beyond, and in addition to, those that already exist at various levels in the school system.

The execution of any decentralising policy of the kind referred to must obviously take some time. Some years have elapsed since it was initiated, and several more will pass before the process is completed.

It will be necessary to establish and define unfamiliar responsibilities for a range of officials — and in some cases new posts will have to be created to cope with these. A variety of new bodies must likewise be established to carry out the tasks, eg consultative councils, education centres with their teams of advisers.

New and established institutions such as district colleges, colleges of education etc, must adjust themselves to new tasks.

Modern western society seems to be approaching what one might call an epoch of "enlightened democracy". Universal education and general affluence offer society opportunities to tackle tasks in new ways and in new contexts. Centralised systems developed such structures as were necessary to permit them to function according to their own lights. Decentralised systems must do likewise. Transition from centrally to locally organised innovation typifies one area of activity calling for re-structuring and delegation.

Unless the political will for decentralisation is backed up by the appropriate executive structure, financial responsibility is delegated, and qualified personnel is appointed locally, there will be no foundation for local initiative and enthusiasm to rest on. Enthusiasm and optimism at local level must be based on a realistic appraisal of conditions, if it is to contribute towards a favourable environment in which decentralised reform can take place.

Decentralised innovation: the planning stage

A common objective for all educational reform is the qualitative improvement of educational opportunity. It may, of course, be questioned whether such an improvement can be achieved at all through decentralisation. A variety of factors will determine the answer to that question, not least the quality and the level of the planning undertaken. One may, nowadays, be permitted to assume that local and regional authorities and administrative bodies are familiar with general principles concerning educational planning. However, while centralised planning by and large takes account of concerns that are common to all regions of the country, local and regional planning can more easily take into consideration needs that may be peculiar to the educational environment of the local community concerned, and even to particular institutions within

that community. It is nevertheless vital to ensure that those responsible for applying general principles to specific situations are properly qualified to do so. If schools are to be run on democratic lines it is obvious that planning — whether central or local — must not be left in the hands of an elite. Experts must be regarded as members of a team, some of whose expertise consists of thorough knowledge of local conditions — needs — concerns — and resources. The value of any proposals or plans drawn up by such a team will thus lie in the fact that they will be specifically tailored for the particular environment in question and in many cases for that environment alone. Freed from the need to take into account matters of general concern, local and regional planning can concentrate its attention in the opposite direction, namely on matters that are of purely local concern.

Centrally sponsored reform has always tended to stifle local efforts. The creation of a nationally uniform school system led to a great number of serious problems; as when attempts were made to establish state schools in various districts, each with its own special characteristics. Most of these problems should be soluble within the context of decentralised reform.

As far as educational research is concerned, decentralisation should offer clear advantages insofar as it will allow for a systematic analysis of existing local conditions, collaboration of shortcomings, and suggestions for possible alternatives. At this stage it is important to consider reports and discussions of completed as well as current development projects, both at home and abroad. The aim here will not be simply to note broad conclusions or findings, but to study the conduct of the various projects in similar environments, in order to discover particular problems and findings related to them, that concern the type of environment in question. Such information is of particular value in decentralised reform, and should provide researchers with intimate knowledge of the local environment from the outset of a project, while at the same time familiarising them with matters relating to procedure.

The development stage: continuous formative evaluation

Reform at local level, aimed at meeting the interests of the locality, calls for continuous evaluation. Continuous formative evaluation will aim at discovering the extent to which such processes as have been established, or put into effect, seem likely to contribute to stated objectives — whether these are of a broad, long-term, or more limited

nature. The objectives of evaluation at this stage are twofold:

- to ensure that what is in fact carried out conforms to stated intentions; and
- to provide an objective basis for subsequent decisions.

The likelihood of achieving both of these objectives will be improved if the research staff involved are themselves associated with regional or local institutions — eg colleges of education, district colleges, educational service centres etc — in this way it will be possible for research staff to ensure continuous appraisal of developments through direct observation, direct discussions and seminars involving other interested parties. It should accordingly be unnecessary to rely on procedures that are more appropriate in situations where contact between those involved in the process of reform and those evaluating it is remote and intermittent.

The influence of research on political and administrative decision making

Decentralised evaluation may be considered primarily as an intelligence service within the organisation of continuous reform (which may incidentally form an actual part of the education service provided by the local community). In this way it is more likely that the contributions made by research staff themselves will keep pace with developments and, furthermore, will be available whenever decisions must be taken. Administrators are often unwilling or unable to devote time to studying lengthy reports or exhaustive documentation. It should be possible for them to become acquainted with the various contributions of research in other ways — eg through discussions focused on those areas within which specific decisions have to be made. If research — development — dissemination — implementation — administration — and policy making — are to be fully co-ordinated, it will be necessary to establish far more effective and fluid lines of communication than has hitherto been the case.

Perhaps the greatest advances are yet to be made at the level of personal contact — which is in any case obviously more feasible in the context of decentralised (as opposed to centralised) reform.

Various kinds of research — development — dissemination — innovation and institutionalisation have perhaps been more satisfactory in social and economic contexts than in education. We are aware of few projects that are directly applicable to actual teaching, or to educational administration.

Headmasters, and principals of schools and colleges have looked in vain for guidance to major research projects dealing, for example, with classroom differentiation. Administrators and authorities with budgetary responsibility have found no clear guidance from research dealing with, for example, optimum class size. It is indisputable that a great number of people associated with the school have found the results of educational research singularly unhelpful.

It is nevertheless important to recognise the potential influence of research findings on political and administrative decision-making in the context of decentralised educational innovation. Such influence may be detected in the general political climate and more specifically in attitudes regarding matters of educational policy, insofar as actual decisions are felt to be founded on rational criteria. Provided that the quality of research can be maintained at a high level, and that the bulk of the research findings can be co-ordinated, it is reasonable to suppose that varied findings will nevertheless display broadly discernable common tendencies. These will, in turn, contribute towards the establishment of the kind of climate mentioned above, and will, of course, help to determine decision-making. Research may have more to offer in terms of general indications, as opposed to specific concrete conclusion. It should not be a function of research to determine political or executive decisions, nor to dictate to, or undermine the authority of executive bodies and administrators.

Education shares with politics a broad concern with life style. Each involves highly complex considerations which cannot be supplanted by educational research on the one hand or political science on the other. Science and research can help to foster a more rational approach without necessarily making the decision-making process any easier.

The follow-up stage

While innovation should form an integrated component of contemporary teaching practice it is desirable to define each new element in such a way as to ensure continuity within the existing framework — and to preserve a clear distinction between temporary expediency and standard practice. Continuous formative evaluation must be complemented by summative evaluation upon completion of any programme of innovation. Summative evaluation should offer a basis for decisions as to whether, and in what form, innovation is to continue: as well as provide pointers to possible improvements in existing practice. The relationship between evaluation and decision-making at

this stage will roughly correspond to that indicated above in respect of earlier stages (planning and development).

In this context, too, it is important for research staff to be in sufficiently close contact with the local community for them to be able to convey research findings to those responsible for executive decisions. Their task at this stage will be not so much to generalise from research findings, as to concentrate on the significance of those findings as far as the reform in question is concerned.

The division of responsibility between centre and periphery

During a period of rapid educational growth and change, it is reasonable to expect curriculum development to be given the "lion's share" of available resources (staffing and financial). This is the case at present. The best work in curriculum development will be solidly founded on relevant learning theory, advanced social psychology, and knowledge of behavioural dynamics — in the school, in the classroom, and in other groups at various age levels.

Not all local environments will be equally well qualified to undertake this kind of work. Moreover, there is a growing recognition of the fact that work on curriculum development in general seems to suffer from weak theoretical foundations, that is to say that fundamental theoretical research has not kept pace with the recent flurry of applied research and the expansion in education generally. The trend towards decentralised reform makes possible a more sensible differentiation between various types of institutions and other bodies. No single institution can claim to be wholly equipped to tackle the myriad of dissimilar tasks that come under the heading of "educational innovation". The universities should be permitted and indeed, encouraged to continue to develop their basic theoretical expertise, while institutions more concerned with practical teaching — eg colleges of education, educational service centres etc — should direct their attention to correspondingly practical tasks. Responsibilities should be allocated to the various administrative bodies in a similar way. There will, of course, always be some institutions and administrative bodies which fall between the two extremes (theoretical and practical) whose proper functions will frequently and appropriately be to ensure contact between the two extremes, to bridge any gaps between the theoretical and practical — as well as between centralised and decentralised innovation.

Centrally directed innovation should be primarily concerned with "macro-level" issues, ie those revolving around such considerations as — "the role of education in society" — "the structure and functioning of the educational system in relation to the goals set for it and the resources allocated to it". Local and regional bodies will of course, take such issues into consideration — albeit within more defined limits — to the extent that they are financially autonomous.

Decisions at "micro-level" — where the emphasis will be on such issues as "the development and characteristics of the individual student" — "practical conditions for efficient learning" — etc, should be taken by local authorities. It should be pointed out here that collaboration between local and central authorities should take place to a greater extent than hitherto.

Between the two extremes there are intermediate level tasks where matters relating to the learning environment, the curriculum and so on are of central importance. It is possible that just such tasks can form the basis of concrete projects in which institutions of various kinds and levels co-operate with each other.

Questions for decision

Important tasks to be tackled during the transition from centralised to decentralised innovation are:

- defining and categorising the specific problems that should be researched;
- attempting to establish or identify the type of institution that is most suitable for coping with the problems;
- deciding the level at which the problems should be approached.

Such procedures should be appropriate and feasible in the context of evaluation in innovation as sketched above.

Research priorities

One example of a plan for research work is that drawn up by the Rand Corporation for the National Institute of Education in the USA. The plan defines four areas in which renewed effort is called for. Each area is further divided into a number of "programme elements". The first of the areas covers such major educational problems as the improvement of educational opportunities for the disadvantaged, improvements in the general quality of education and the achievement of greater efficiency in the use of educational resources.

The second area aims at improving educational practice while the third aims at strengthening the foundations of education — selective research programmes building basic knowledge concerning education. The fourth area is entitled "Strengthening the research and development system". A further step would be to work out a plan of what institutions and levels should have the main responsibility for the different areas or programme elements.

Another example of an attempt to establish priorities for research tasks may be found in the Norwegian Department of Education's Statement of Intent regarding educational research. The department has suggested that research should be carried out in the following areas from 1975 onwards:

- the ways in which the new Norwegian outline curriculum ("Mønsterplan") is being put into effect in primary and lower secondary schools,
- integration of handicapped children into the general school system,
- developments in schools at upper secondary level,
- the place of the school in the local community,
- analysis of the effects of decentralised executive authority within the education system,
- evaluation of high cost teaching materials,
- investigation of the relationship between special education and opportunities for the physically and mentally handicapped in society,
- the roles of the sexes in education,
- identification of groups availing themselves of opportunities in adult education.

It is open to any institution, group or individual to apply to the Department of Education for full or partial financial support for projects within these areas. The Department of Education itself

has taken active steps to draw local institutions into this work, illustrating one way in which central authorities can tackle problems in which they themselves are interested, at the same time as they stimulate local interest in research evaluation. But so far we have derived little experience from the involvement of various institutions at different levels in solving problems within broadly defined research areas.

The Norwegian Department of Education also supports research in other ways, eg through grants to central organisations such as the Norwegian Council for Innovation in Education, the Norwegian Research Council for Science and the Humanities, as well as local bodies. At present, projects initiated locally must, to a large extent, rely on local financial resources. However, working parties on educational research are now preparing proposals for co-ordinated programmes which will incorporate plans for financial support for educational research carried out at all levels — local, regional and central.

Conclusion

It appears obvious that problems associated with the trend towards decentralised innovation and evaluation are not simply financial, organisational or administrative. Neither decentralised innovation nor evaluation are likely to be wholly successful unless the institutions and individuals involved (research workers, teachers, administrators or local people) are given broad institutional support, as well as theoretical and practical training. At the same time it is to be hoped that improvements in the quality of education provided in each locality will contribute towards the establishment of a political climate which will in turn ensure continued development.

Local variations in national curriculum development projects: self-generated school-based projects and their implications for evaluation

by F. H. SPARROW,
Schools Council, London.

I want to begin this talk in a slightly unorthodox way by referring to a minor error in the documentation. You might have noticed that one document refers to background information from the "U.K.",

which means, of course, the "United Kingdom". If any of you have followed the events of last week, particularly the General Election last Thursday, you might consider that the adjective "United" is

open to some question. If this constitutes an "error", it is not the one to which I refer. The information comes from England and Wales and I must make it clear that I will speak only of these. One of the things which England has in common with Scotland is a different educational system and I have no brief or competence to speak for Scotland. In my talk I will probably refer only to England, but this will be purely for the sake of brevity and is not to be taken as a gratuitous insult to Plaid Cymru, the Welsh Nationalist Party which won one notable victory over the British Labour Party last week.

This talk covers six main areas.

- A brief resumé of the English system in the light of the title of this symposium.
- Some illustrations of how curriculum development grows from and is influenced by this system.
- Some aspects of evaluation trends.
- Reasons why the Schools Council is currently considering support for self-generated, school-based projects.
- Some examples of local projects which are being considered for support.
- Possibilities for future development and their implications for evaluation.

The English system

The essential point which has to be stressed is that it is de-centralised. The Department of Education and Science is the official Government Department and it has a great deal of power. It administers the law and it naturally plays a leading role in the evolution of new laws and new policies. It has considerable control over funds, a control which is, perhaps, most apparent in respect of school buildings. Her Majesty's Inspectors act as the eyes and ears of the Department.

We also have over 100 Local Education Authorities, a fact which surprises many Germans, with their eleven Länder and Swiss, with their 22 Cantons. Indeed, until last April we had over 160 of these Authorities. Each has a great deal of power. It is responsible for implementing the law within its own area. It handles the money which comes from local sources in the form of rates. It submits schemes for Comprehensive Schooling to the Department and in this connection it is interesting to note the extensive range which exists. We have Comprehensive Schools for the ages 11 to 18, 11 to 16 (followed by open-entry Sixth Form Colleges), 12

or 13 to 16 or 18 (where there are Middle Schools for children up to the lower age) and so on. Indeed the confusion over terminology to which Professor Mitter referred is very apparent in England. The Local Authority maintains schools, with their buildings, equipment and staffs. It employs Local Inspectors who are commonly called "Advisers" these days.

We also have thousands of schools, each of which has a great deal of power. The most powerful single figure is the Head Teacher, who has been referred to as "the last autocrat in England". Nowadays there is an increasing trend towards a hierarchic structure of Deputy Heads, Year or House Tutors and Heads of Departments. Even so, the most important person in the last resort is the classroom teacher and he too has a great deal of autonomy. In effect, nobody, whether from Department, Local Authority or School Office can tell him what to teach or how to teach it. In case you think that this must be a myth, let me illustrate from my own career. I taught in England for twenty years in many different types of school. In all this time I saw Her Majesty's Inspectors on only six occasions, I was never visited by any Local Authority Adviser and no Head Teacher ever heard me teach. It is a fact that nobody ever told me what to teach or how to teach it.

From this it will be clear to you that if we are talking about a general trend towards de-centralisation England is already more than half-way there. However, I must add a caveat. There are in fact many influences at work which give the system more cohesion than appears at first sight. English teachers come from what might be described as the same general milieu. Although there are certainly wide regional variations there is also a lot in common in the educational ethos and background. Moreover, many teachers experienced a similar schooling, some aspect at least of which they perpetuate in their own teaching. At its worst, when a teacher has been educated in a very traditional school, followed by a formal University, this can lead to what has been called "the cycle of stagnation". Another influence is the frequent meetings of teachers, ranging from the informal to Courses and Conferences arranged by many possible agencies, including HMI, LEA Advisers, Professional Associations, Subjects Associations, University and Colleges and Examination Boards. English teachers work less in isolation than ever they did. Another centralising influence is the external examination system, about which a lot could be said in another context. The situation is, of course, much more complex than so brief an account can possibly reveal, but my intention is to demonstrate

that curriculum development can take place in England only if it is undertaken in ways which neither threaten nor challenge that autonomy which is sacrosanct.

Illustrations of English curriculum development

In the early 1960's the Department of Education and Science made a move to stimulate curriculum development by setting up a central Curriculum Study Group. The result of this was an outcry from the teaching profession. It appeared to them that their cherished autonomy was about to be overthrown and that a central authority would soon dictate the curriculum to the schools. However true this was — and I, as a teacher, was certainly convinced of its truth — the teachers made it clear that they would have no part of it. The result of this was eventually the setting up of the Schools Council, the organisation which I represent here today. Most of the curriculum innovation initiated in England since 1964 has been under its auspices. It is a central organisation which works in a de-centralised and non-directive manner. The point can be illustrated by reference to various aspects of Council Projects.

The decision whether or not to sponsor projects is made by policy Committees, all of which must, under the constitution, have a majority of serving teachers. In this way it is the teachers who decide what is to be undertaken. The projects are based on centres spread almost throughout the land, so that although the sponsorship is central, the projects are widely dispersed. Project teams normally include serving teachers who are released from their teaching duties for the relevant period of time. Methods and materials are tried out in schools which are chosen not for their assumed excellence but for the way they represent schools as a whole. The schools are approached only through their Local Authorities. No school and no teacher is obliged to co-operate in project work. Teachers who take part are intended not to be merely passive receivers of packages, but active participants in a common enterprise. Methods and materials are never intended to be definitive or prescriptive. It is assumed not only that teachers will feel free to select material and add their own, using methods which vary according to their perception of their children's needs, but also that this teacher freedom is good in itself. When projects have completed their trial stage and published reports, teachers' guides and pupil material are available, schools are completely free to accept or reject what is offered. The Council's aim is not to impose curriculum development on anyone, but rather to widen the range of choice available to teachers.

There is always a danger that speakers in my present position will rationalise the status quo; that having to accept situations which they cannot control, they will make the theory sound good, no matter what the reality of the practice. In case you think that what I have said about projects reveals self-delusion I can assure you that if we took that long list of projects which you have been given, I could show you how each one illustrates what I have said. In the limited time available I can only use examples. One of these projects, *children as readers*, has no defined central programme. It stimulates groups of teachers to conduct their own development work. The *history, geography and social science* project starts from a number of core concepts, but it does not intend to produce materials for all teachers so much as to assist teachers to develop their own, tailored to the needs of their children in their situation. The *continuing maths* project produces no pupil materials centrally, but a large quantity of material is produced by locally-based groups of teachers, with help and support from the central team. The materials produced centrally by the *humanities curriculum* project have sold well, but it is estimated that its users fall into three categories, one of which applies the methodology of the project as closely as possible, another uses the material as invaluable support for teaching in ways which owe nothing to the project, while a third probably mixes the project philosophy with various preferred teaching styles.

Some aspects of evaluation trends

It follows from all I have said up to this point that the evaluation of curriculum development in England cannot possibly be a simple process; indeed, as each evaluation is geared to a project rather than to some abstract external model, no two of our evaluations are alike. However, after over ten years' experience it is possible to detect certain trends which reflect the decentralised English system. For example, Her Majesty's Inspectors were used extensively in early evaluation. This was in accord with a long-standing tradition. Although it can reasonably be argued that such evaluation is subjective, it must also be remembered that these Inspectors were in a good position to compare schools and that their pooled information represented more than a narrow personal judgement. In any case, it is probably wise occasionally to remind ourselves that evaluation means "valuing" and that in the last resort this must be largely personal, no matter how objective the evidence may appear to be. When the practice grew of appointing evaluators or evaluation teams to Schools Council projects there was a very strong tendency for these

first "professional" evaluators to be influenced by the engineering model. The stress tended to be on measurement: pre- and post-tests, control groups and the like. In the last four or five years there has been a distinct movement away from this approach. There are many reasons for this. The English scene which I outlined earlier is hardly amenable to exclusively measurement procedures. There is no centrally determined curriculum, the various levels of autonomy I described have to be respected, there is no attempt whatsoever to produce "teacher-proof" materials and the teachers themselves are closely involved in the development process. In such a context the stress has had to be increasingly on local variations rather than on national measurement. Perhaps England is not as unique as I appear to be suggesting and that what Dr. Wynne Harlen has recently written for a book we are hoping to publish next summer will find an echo wherever there is curriculum development.

"When put into practice this classical model, which seems logically attractive, fails to give the information which is most useful either for formative or summative evaluation. What makes it so attractive and reasonable is precisely: its weakness; it is part of an oversimplified view of curriculum development as 'identification of objectives, planning of learning experiences, evaluation' (Wiseman and Pidgeon, 1970), which ignores the characteristics of the learning environment and the interaction of learners and teachers with it, both in the development and evaluation stages. In reality what pupils learn in any situation depends on a complex collection of factors which have to be taken into account if the evaluation information is to be of practical value."

It is certainly the case that evaluation in England has concentrated increasingly on "process" rather than "product". We want to know not so much the estimated total gain but more the context in which certain results are found. The number of variables which affects these outcomes is very large indeed. Our evaluators collect as much data as possible about them, using a wide range of the techniques available. In addition to organising feed-back from children and teachers, some of them use tests of attitude (tending to use scales developed for the purpose as existing ones are rarely suitable), observation and interaction schedules (again with a tendency to develop original scales) and an increasing number follow a descriptive, case-study approach. Dr. Harlen has used an interesting technique for pulling together those variables which are amenable to coding. It is a form of cluster analysis and she describes it in her evaluation of Science 5-13, which will be published early in 1975. The

most clearly defined intention of all this work is to provide information for decision makers, whether they be a project team, classroom teachers, Headteachers, Local Authorities or whatever. It is interesting to note that this attitude to evaluation in England existed long before the work of such writers as Robert Stake or Parlett and Hamilton was well-known in my country. Equally, of course, the Research Team of the Schools Council, when asked for advice by newly appointed evaluators, always suggest that they should read as much as possible about evaluation theory from all sources. It is important that they select from the possibilities those aspects which are suited to their particular needs. Evaluation is probably most irresponsible when those concerned with it allow the apparent brilliance of one school of thought to dampen their own critical faculties. Just as English teachers are encouraged to choose from the curricular opportunities which are available, so evaluators are urged to choose from a wide range of possible evaluation methodologies.

Reasons for the beginning of a change in council policy towards self-generated, school-based projects

About a year ago, the main policy Committee of the School Council decided to begin to explore ways in which local curriculum development might be both stimulated and helped more directly. No doubt there were many reasons for this proposed change of policy and no formal statement of motivation was made. However, it seems to me that there were two strong reasons, one partly negative, the other strikingly positive.

The negative reason is that we, in common, I understand, with most other countries which have attempted curriculum change, have become somewhat disillusioned with the R, D and D model. The chief problem arises from the final "D" — "diffusion". The plain fact is that, no matter how good a curriculum development project might be, dissemination, diffusion and take-up have tended to be disappointing. Equally, if we think of the "Centre-Periphery" model, the work at the centre might have been good, but the spread outwards to the periphery has not always taken place. It was for this reason that the Schools Council set up a Working Party on Dissemination, which has recently produced a report and recommendations for action. We have to remember that there are many forces working against curriculum development. In England, teachers have been thrown into tremendous changes in school organisation at the very moment in time when bodies such as the Schools Council were urging changes in the teaching and learning process. Many teachers have been so preoccupied

with the change to comprehensive schools that they could reasonably claim that other changes are inappropriate or even impossible at this moment. Similarly, the rapid nature of change in Society has imposed heavy burdens on teachers, many of whom are so preoccupied with their day-to-day problems that they have little or no time or inclination to take on the additional demands which curriculum development is likely to make of them. There is a paradox in this situation in that much of the development which they might otherwise take up is aimed at helping them to solve some at least of their immediate problems. Something clearly has to be done about this. One possible solution is to put some effort into looking more closely at the "Periphery-Periphery" model. The theory behind this is that teachers might be more amenable to precept than to preaching. If they can see at first hand, perhaps in their own schools, perhaps in those of Teacher Centre colleagues, the advantages which could accrue to them from participation in curriculum development might become more apparent. There are, of course, many difficulties in this "P-P" model and it would be most unwise to put it forward as a panacea, but the fact remains that there is a place for more initiative at the periphery and somewhat less at the centre.

The positive reason is simply that although there is a tendency to think of curriculum development as a feature of the 1960s, it has, in fact, been a continuing process over time. There have always been teachers who have tried out new strategies, new methods of teaching and who have produced their own materials of various kinds. It can be argued that all that is new is the elevation of curriculum development to a more overtly important and sponsored position. This is why many of the best projects have been largely the work of successful practising teachers. However, nobody would pretend that anything like the total potential resources have been tapped. If ways can be found of locating the best work which is going on, stimulating it by practicable means and making its contribution more widely known and more generally available there could well be a new stimulus to curriculum change and one which, because it is firmly rooted in the reality of teacher needs, could be more effective than the traditional C - P model.

Some examples of local projects which are being considered for support

When this intention of the Schools Council became known there was an immediate response. Many groups of teachers, together with College of Education and University staffs began to seek support

for local projects which were already in existence but at varying stages of development. I will give a few examples by way of illustration. One of our Local Authorities has been working for 5 years on a Mathematics project which aims at providing support, guidance and teacher and pupil material. The demand for this help has been so strong that the Project Team have had no opportunity properly to evaluate what has already been done. They have therefore asked for support for a much-needed evaluation programme. A group of teachers in another area has been looking at particular points of difficulty met by pupils in Mathematics. They have asked for help in this work and its extension to methods by which teachers deal with the situation. Two Colleges of Education and one University Department have been working with local teacher groups on Social Science projects and they too have asked for support. Similarly a number of teacher groups have been concerning themselves with the problems of mixed ability teaching, and they too are seeking support. One of the most interesting of these local schemes is one which comes from the existing work of a group of teachers at a Teachers' Centre. They are concerned with problems of communication, particularly oral English. It is interesting to note in this connection that the teachers concerned have made themselves familiar with recent research into communication. Perhaps the gap between researchers and teachers is not as wide as some people have suggested.

The point which is common to all these requests is that they are all, in some degree, local equivalents of national projects. In so far as this true it will be clear to you that evaluation strategies will have much in common with current practice in national projects. However, there are two reservations; it will clearly be much more difficult and dangerous to generalise from the "local" than from the "national"; evaluators will also have to be continuously conscious of the fact that experience in the process of curriculum development is likely to be much more important than the product.

Possibilities for future developments and their implications for evaluation

You will have noticed that in the list which I have just given to you there is nothing from individual teachers. This is natural in that groups already in existence obviously start from a position of advantage. There are many problems in stimulating teachers to seek support for their own curriculum development work. At this moment of time the Schools Council has not evolved procedures for locating interesting work, selecting the most appropriate and solving the problems of rela-

tionships with the Local Authorities concerned. It could be that most of the support available will be given to groups of teachers rather than to individuals. However, if proposals do come in from individual teachers and schools, and if they attract support, a major problem, and one which is currently exercising the Schools Council Research Team, will be how such schemes are to be evaluated. We do not yet have a positive answer to this question, but three points are fairly clear. The first is that the evaluation procedures followed in national projects will be largely inappropriate. It would be impossible to attach a "professional" evaluator in each case, and great care will have to be taken to ensure that evaluation does not in any way inhibit the natural development from which the work sprang. Secondly it is clear to us that in this new situation the evaluator's role will be that of guide, counsellor and friend. Thirdly it seems certain that in work of this kind the teacher concerned will have to be, to some extent at least, his own evaluator. The dangers of this will be apparent to you all, but the

fact remains that in very small school-based projects it is only the teacher who can monitor his own efforts, clarify his intentions and estimate his level of success. We should remember that good teachers are continuously evaluating both their pupils and themselves. They need help if their judgements are to be more than merely subjective.

Conclusions

I gather from conversations with the people present here that although the contexts in which we all work are different many of the fundamental problems are essentially the same. We have no glib or facile answers, but the first problem is always adequately to define the problems. One important way in which progress can be made is the exchange of information and experience. Such an exchange is facilitated by this conference and I am grateful to the Council of Europe and to the Federal Republic of Germany for making it possible.

Teachers' behaviour as a variable in evaluating school reform pilot projects

by L. LEGRAND,
Research Director, Institut National de Recherche
et de Documentation Pédagogiques, Paris.

One of the most formidable problems facing educational authorities in their search for improvement is to get teachers to accept the reforms they have decided upon. We should note and deplore the fact that most of the time innovative decisions are made with no preliminary estimate of the conditions necessary for success or of the inevitable cost of achieving these conditions.

The responsibility for this state of affairs rests most often with over-hasty and voluntaristic decision-makers, but it very often rests as well with the researchers and innovators who do not pay enough attention to the human factors involved in developing the innovation judged desirable or indispensable. Most of the time indeed innovations are concerned — consciously and principally — with curriculum content, teaching methods or educational structures; and the impression is given that these are regarded as the only important things in the innovation process. However, in the early stage of any innovation a change in the teacher's attitude and behaviour is often more important for the success of an innovation than

are changes in curriculum content, method or structure. Even though the Hatorne effect is well known, it is rarely taken into consideration in decisions to introduce innovation on a general scale. This is perhaps because the mechanism of this effect has not been mastered precisely enough to give technical instructions on how to produce it.

The main aim of my paper will be to highlight a methodological problem: how, when an innovation is being evaluated, can the teacher's behaviour be taken into account so that the educational authorities can allow for it when taking decisions on the introduction of the innovation on a general scale?

The importance of teachers' behaviour in any innovation

It is now standard practice to distinguish between three types of innovations:

1. Innovation which spreads from the centre to the periphery

A new content, a new method, or a new school structure is produced and "laboratory" tested. In the light of an evaluation of the results, a decision to develop the innovation is taken and, after further evaluation, this is followed by a decision to apply the innovation on a general scale. This is the standard innovative process in centralised countries. Moreover, the intermediary stage is often "forgotten" for lack of funds and in a desire for speed and pseudo-efficiency. This process fails quite often, the innovation usually being rejected by the teachers or at best deformed, eroded, neutralised. Consideration of this failure classically reveals the importance of the teacher's behaviour and of the error mentioned above of not taking his behaviour into account in decisions to generalise.

2. The second type *spreads from one peripheral point to another*

This is diffuse innovation, produced by teachers on the spot themselves, and spreads from one point to the next, somewhat like a fashion. The behaviour of teachers as creators and partisans of an idea here is of prime importance. But the process entails no evaluation. Either it spreads rapidly often disappearing as quickly as it appears, or it isolates itself in pockets.

3. The third type seeks to combine the advantages of studied and evaluated innovation with those of spontaneous innovation. With this type of innovation, a dialectic is *established between the centre and the periphery*, and during this dialectic there are phases of routine evaluation. The creative initiative belongs to local teams who have decided to join together to solve a common problem. The centre comes in as a catalyst and instigator more than as a creator. Generalisation is linked to the progressive creation of curriculum content, method or structure; and evaluations intervene as a regulating force in the creative process. Here too, it rapidly becomes apparent that the teacher's behaviour is a fundamental factor in the success or failure of innovation. Whatever sort of innovation is under consideration, though, the job of evaluation, if it is to be effective, should be given to a team distinct from the teaching teams. The ties which exist among these teams, however, and also the place of teachers in the evaluation mechanism itself are both decisive factors for the validity of the evaluation.

Taking the teacher into account in the evaluation of conventional innovation (from the centre to the periphery)

In this case evaluation is carried out at two stages in the process. *First of all at the conception stage.* In the case of producing teaching material, evalua-

tive checks may be numerous, following each sequence produced. In the case of structural innovations, they are less frequent and occur from time to time when each decision has been followed by a period of application (six months, one year). At the ideas stage, the teacher's behaviour is very rarely taken into account. The important thing for the planning team is translating its objectives into terms of pupil behaviour (tests, questionnaires). The teacher in the school situation is then a member of the planning team and is so de facto in a continuous training and maximum motivation situation. He is rarely the subject of observation. When he is, it is part of keeping a check on the operation, the evaluator being an integral part of the planning team and paying more attention to the pupils at grips with the material taught than to the volunteer teacher chosen for the experiment from those considered a priori to be good teachers.

On the other hand there is considerable concern with the teacher's behaviour fully, at the development stage (full-scale trial of the innovation product). There are two possible attitudes here.

The more usual, and the easier, is to try to neutralise this factor through recourse to criteria extraneous to the actual act of teaching. The evaluation mechanism consists then, in addition to information gathered about pupils (general tests, sociological data, tests putting the innovators' objectives into practice), of information about teachers carrying out the innovations. This might include sociological files (age, sex, type of training, attitudes, interests) so as either to neutralise the teacher factor (observation of performances of comparable pupils taught by comparable teachers) or to illustrate the possible influence on comparable pupils of the variables in the teacher files. This might be called a macroscopic evaluation of the teachers' behaviour. Such an evaluation method is useful when a rough, large-scale judgment of the effectiveness of an educational system is needed, but it is clear that it does not capture the essential.

It is noteworthy, indeed, that a closer study of the performances of comparable pupils, taught by teachers who seem comparable according to the sociological data assembled in the teacher file, reveals group phenomena (performance of classes taught by a particular teacher, or performances of a school taught by a group of teachers) which remain quite unexplained at this level of analysis. From the point of view of the effectiveness of an innovation this subtler approach might well be the key one. A closer look should therefore be taken, and the reasons for a better or worse performance investigated. Teachers' behaviour should be described in an objective and, if possible, quantitative

manner to equate this behaviour with the other variables in the system studied. The various methods of objectively analysing teachers' behaviour are essential here. They are becoming well known through the works of Flanders and, in French, those of De Landsheere and Chobeaux. These analyses, involving the analysis of features of verbal behaviour are valuable for shedding light on periodic phenomena. They will nevertheless have to be improved to include non-verbal factors and fit types of non-frontal teaching (work groups, "animation") where are now difficult to handle.

Two methodological remarks about this type of study are appropriate here:

To begin with, their large-scale use is impossible owing to the number of observers necessary, and to the need to train them to ensure methodological agreement. For full-scale evaluation, the only dependable source of information for decisions, it is absolutely necessary to work with enough pupils to have a representative sample of the national school population as regards sex, age, general development, and socio-professional background. Thus the number of classes and teachers involved should itself be fairly large especially if all these teachers are meant to provide a good sample of the teaching body concerned. To pinpoint the variable which teacher behaviour constitutes, a more manageable instrument must be used. There is no evading the difficulty: only a questionnaire could meet this need. But the questionnaire must be based on detailed behavioural studies and deal with clearly defined forms of behaviour. The major obstacle is the discrepancy between what the teacher thinks he is doing and says he is doing, and what he actually is doing. The questionnaire must therefore deal with facts and not opinions (eg: "How many words did you make your pupils learn by heart during the week beginning...?", and not "Do you make your pupils learn texts by heart often, sometimes, not at all?"). It must also be filled in by an investigator and supported by proof.

The second danger is that teachers' behaviour is liable to be studied in the light of certain standards. It is the natural temptation when the tested product has been previously developed in the laboratory and is now being "sold" with instructions for use. There is good and bad use of material. This is the case with pupil tests who have been evolved in the laboratory and become a standard of reference for the evaluator with evaluation in the form of measurement of deviation from the norm. The failure of the pupils can then be ascribed just as easily to the behaviour of the teacher as to his misinterpretation (from the evaluator's point of view) of the objectives implicit in the products used.

It is tempting in these circumstances to blame failure on inadequate training, which may be the case. But it may also be due to the basic inadequacy of the new material in real school situations, resulting from an inefficient or non-representative number of teachers and pupils having been involved in planning the innovation. An imposed objective is generally doomed unless account has been taken, from the beginning, of the real state of the educational system. New curriculum content in mathematics, for example, needs to have been created from the beginning, by involving all and sundry, teachers and pupils, in the process, if the result is not to be a merciless instrument of social selection, as has often happened with this subject.

This is why, when an innovation has come from the central authority it appears advisable to observe the educational system and teachers' behaviour just as they are, without any preconceived ideas, an attempt being made to work out objective typologies related to pupils' performances, which should themselves have been observed without reference to any pre-determined standards. Such action may certainly hold surprises for the central policy-makers on innovation.

Taking the teacher into account in the evaluation of supervised innovation (from one peripheral point to another via the central)

In the preceding descriptions I have taken for granted the co-operation of teachers, the subjects of evaluation, and this is an aspect that must not be overlooked. Evaluation at the development stage of a downwards innovation always encounters a more or less clearly-stated refusal by teachers to be judged from the outside as operatives and subjects for study. Teachers habitually challenge the validity of pupil-tests in such operations. This not only reflects emotional reactions, but often points out differences between the objectives of the "policy-makers" and those of the experimenters. In the same manner the evaluator's surreptitious recourse to the norm is always seen by the teacher as a personal attack.

It is clear that the evaluation of spontaneous innovations would be viewed as being even more oppressive.

This being the case, innovators almost always reject outside evaluators as being inadequate. When evaluators are imposed upon them, the teachers receive them distrustfully and challenge their methods and findings, calling them irrelevant to the objectives sought. This is frequently the case, moreover, as these spontaneous innovations usually result from ideological and ethical considerations

that are very difficult to express in any final, observable behaviour, if only because of the long-term effect assumed by the innovator. But it is too convenient for the innovator scornfully to reject anything he cannot catch in the nets of an objective methodology as "pipe-dreams".

Thus, the evaluation of a spontaneous innovation has no real chance of success or effectiveness unless it is not only accepted by the innovators but is actually requested by them as a means of facilitating and regulating the innovation. Terminal (or summative) evaluation can only result from a succession of partial, regulatory evaluations in which the teachers-innovators have an important role to play.

There are two cases where the conditions necessary for such acceptance and participation may exist. Either, the innovators themselves will ask for the support and criticism of an evaluation team. This presupposes much mutual confidence and usually a lowering of the initial level of enthusiasm found with any new spontaneous innovation. There must be at least concern or suspicion about the real effectiveness of the action undertaken.

Or, the idea will come from the centre, albeit in a general context favourable to the idea, which is merely officialised by the centre. In this case periphery needs to be appealed to for a voluntary commitment to an innovative action, in which each of the participants has from the beginning an equal voice in the discussions and experiments, and consensus is the result of continuous product comparison and evaluation.

Such a process enables the distortions and bottlenecks of conventional innovation processes previously described (from the centre to the periphery) to be avoided. But several conditions are necessary:

1. The teaching teams must be involved in the designing of the innovation from the very beginning.
2. The pupils and teachers taking part must form a representative sample of all those who would be affected by the introduction of the innovations on a general scale. In other words, the teachers involved should not be hand-picked for their merits or ability; their readiness to participate should be the only consideration.
3. The teaching teams should be organised in the form of school units, each of which should be taken charge of locally by a design and evaluation team, organically and democratically linked with the co-ordinating central authority.
4. The various teams thus organised should be involved not only in the innovation itself (method, structure), through local trials and periodic com-

parisons of results, but also in the progressive evaluation of the innovation, the final instruments being a synthesis of the partial instruments devised by them as part of the innovation.

5. The teachers should themselves be involved in the conduct, marking and interpretation of pupils' tests.

As the research will immediately become full scale, the teachers' behaviour will accordingly have to be taken into account from the beginning of the innovation and not only during the development phase. There will have to be investigations, from the beginning, into how the teams organise or fail to organise themselves, how they interpret the general initial objectives and the theoretical information given, how they apply them, the positive or negative role of the group leaders, the scale and nature of co-ordination, the possible influence of local material conditions. Only if this is done can an informed decision to introduce an innovation on a general scale be taken by the education authorities as the general application of an innovation must recreate the innovation processes as training processes for the teachers.

Teacher observation techniques will be those previously indicated for conventional innovations, namely, questionnaires based on and verified by firsthand observations.

Clearly, the cost is much higher than with conventional operations — too high, it is tempting to think.

It seems to me, however, that the following considerations make the process not only desirable, but indispensable to anyone who really wants to change the educational system and not just its statutory appearance.

1. Such a process can only be engaged in when there is an adequate political consensus on the need for change. The failure of a structure, inadequate pupil performance, a change in the way a subject is studied at university, all provide such occasions.
2. Full-scale innovation must cover — from the beginning and on a voluntary basis — the largest possible number of training centres since a supervised innovation process is already one of continuing education. It prepares innovating centres to play locally, in the phase of general application, the role of encouragement and co-ordination played by the national centre during the research phase.
3. Such a process, if it is to stem from a national consensus on the problem posed and from the need for a solution, must in no way be the product of

preconceived notions regarding the nature of the solutions to be applied. These solutions must emerge progressively from encounters in which the various parties concerned (teachers, parents, pupils, local representatives) participate throughout. In

such a process of change, observation of teachers' behaviour becomes a dialectical moment, and, at the same time, a means of becoming scientifically aware of the conditions necessary for the success of a generally applied innovation.

Evaluation as a system component of educational development

by Professor G. PETRI,
Zentrum für Schulversuche und Schulentwicklung,
Graz.

It has often been observed that there are gross contrasts between the high pretensions of attempts at innovation in the area of education and the modest results obtained. The reasons for these contrasts most likely lie in an under-estimation of the complexity of the problems, as well as in the rush towards and pressure for spectacular solutions. Such solutions are often more influenced by the interplay of special interests as well as local determinants leading to narrow views of the matter at hand rather than by an objective and comprehensive research approach.

How can educational development be made more effective?

Empirical educational research is still in the early stages of its development. On the one hand, the manner of working and the expectations are, in many respects, still under the influences of the speculative pedagogical tradition which often explains things very simply. On the other hand, its methods have been borrowed from the natural sciences and technology, which deal with more easily defined, clearer and less complex structures. Therefore, it may be assumed that the hitherto unsatisfactory innovational techniques — which, in many cases, are based on an oversimplification of the problems involved — could be developed to a greater degree of efficacy if one were to systematically accentuate and show regard for the complex character of educational matters.

Everyone involved in this task needs to go through a new learning process: pupils, parents, teachers, administrators, various interest groups, politicians, researchers and experts in the field. In order to be able to cope with such a difficult learning task, special methods and organisational structures are necessary.

The *methodology* involved in the initiation and permanent establishment of an efficient process of

innovation is a kind of *systems analysis*. This would lead to more intensive consideration of evaluation standpoints as related to all phases and components of the developmental projects.

In the *organisation* of the developmental activities, it is necessary that there be *systematically co-ordinated co-operation between practical experience, development and research within the framework of appropriately selected key projects* on a national and international basis. This necessity follows from the systems analysis approach to the problem.

During the last few years, the meagreness of the results of numerous innovations has become increasingly apparent as a consequence of technically better and more intensive evaluative endeavours. In reaction to this, many new undertakings are in the making for co-operative, comprehensively evaluated developmental activities. These are also a by-product of more dynamic educational policies. It is important now to further develop and to broaden these early attempts.

A systems analysis approach

Systems analysis, as used here in this connection, appears at first glance to be an extremely trivial heuristic strategy to solve complex problems: real or conceptual facts are considered components of a comprehensive system within which these facts are interrelated in a particular way. Elements of a system can again be looked upon as systems themselves; and systems can be examined as components of more comprehensive systems.

In the treatment of problems by means of systems analysis one aims at determining, as *completely as possible*, the systems relevant to the specific problem and their internal and external relationships. Whenever the solution of a problem is facilitated by mapping procedures, then graphs or

mathematical models are constructed, or classification schemes for ordering goals, actual facts, working concepts, alternative hypotheses, open questions, etc are worked out. In this way, the systems analysis used initially under rather general aspects gradually yields to techniques of analysis specific to the particular field examined.

From developmental projects planned and directed by a systems analysis methodology better results can obviously be expected than from an improvised procedure considering a smaller number of relevant factors in a less systematic manner.

Various problems arising in this connection can only be briefly touched upon here. Above all, responses are sought to the following questions:

- To what extent is the systems analysis method a learnable heuristic technique?
- To what extent can it be formalised and developed?
- To what extent is thinking in terms of systems analysis a normal consequence of research and development in the various scientific and technical disciplines?
- To what extent is it based on relatively stable characteristics of personality and intelligence that cannot be easily influenced?
- How efficient is the systems analysis method with respect to various types of projects?
- What forms of negative side effects could appear in connection with the systems analysis method? One could imagine, for example, that the treatment of the problems at hand would not always be aided by extensive use of graphs or mathematical models or by classification systems. Models and diagrams can possibly impede or fixate thinking.

A first step towards the practical consideration of aspects of systems analysis would be the *comprehensive evaluation* of developmental projects. This evaluation would not only relate to the desired results, but would necessarily need to accompany and control all the components of the conceived project, viewed as a system: setting of goals, planning, working out, implementation, and intentional and unintentional effects.

Such a comprehensive evaluation should bring developmental projects into an integral relationship with the systems of educational goal setting, the fundamental findings in psychology and sociology, teaching methods and instructional technology, as well as and in particular, with learning and teaching as practiced.

Reform of project organisation strategies

An *organisational reform of research and developmental activities by means of an extensive participation of teachers on the job*, would make it possible to realise not only greater democratisation, but also:

- the necessary intensification of action research;
- the development of efficient implementation techniques;
- the gaining of teachers, who have proven their abilities in developmental work, as qualified staff members for the centres;
- and a theoretically well-grounded and at the same time action-related system of teacher training.

Examples

Two types of project seem particularly suitable for preparing the way for a co-operative developmental system which appropriately considers the evaluation standpoint. These projects are:

- experiments in changing the school structure;
- projects related to providing for continuous feedback of learning achievement to pupils and teachers as an essential diagnostic, motivating and instructional function.

Two projects have been selected from the programme of Austrian school experiments to illustrate these points:

The structure-changing project "*Experiments in the school for the ten to fourteen year olds*" is mainly concerned with the development and evaluation of a form of *comprehensive school* in which mother tongue, English and mathematics are taught in achievement levels (setting). All other subjects are taught in heterogeneous classes. A political consensus could not be reached on the introduction of a comprehensive school system. The political parties, however, did agree to a law which prescribed evaluative experiments along these lines in order to gain information necessary for further decisions. Several details of this project are contained in the background paper, Part I. In the project under consideration, the following points should be stressed:

- The comprehensive school project raised among politicians, administrators and teachers a marked demand for evaluation and initiated a learning process related to an understanding of the principles of evaluation and the possibilities it offers.

- An institutionalised co-operation was established between university research, school administration, and the Centre for School Experiments within the framework of a scientific advisory board for the evaluation of school experiments.
- In connection with the teaching of synchronised curricula for three different levels in the respective achievement groups, many teachers developed a pronounced problem-consciousness in the selection and structuring of learning objectives as well as in assessment procedures. Teacher groups work together with educational researchers and experts in attempts at setting up curricula.
- Forms of instruction are being set up which systematically try to take into consideration differences in the pupils' learning behaviour: learning objectives are being adopted to the needs of the various achievement groups and individualised learning materials are being developed.
- The problems of the low achiever, who comes chiefly from the lower social levels, became clearly visible to the teacher. The teachers recognise that these pupils require special help so that they are not continually frustrated by failure. This insight emerges in particular as a result of more exact measurements of performance in the experimental schools; as well as in connection with the introduction of English instruction which seems to be very difficult for this group of pupils.
- Finally, the comprehensive school project has succeeded in triggering off a lively and continuing discussion between politicians, representatives of interest groups, teachers, administrators and scholars, on educational problems, which are based partly on provisional results of evaluation.

The second project to be described here relates to *formative achievement testing and the optimisation of psychological learning conditions*. It carries the working title "Learning Test Project". At present this project is in the stage of development of materials. Part II of the background paper contains information about several of its fundamental principles.

The learning tests offer prompt and detailed feedback of the learning results to the student and to the teacher. On the basis of a simple presentation of data, insight into long-term learning development can also be gained. Learning tests differ from the usual formative tests in that they do not fulfil only diagnostic functions, but, on the basis of their specific structure and application, also serve as

learning material. When the pupil goes through the various test items, he goes through the process of problem-solving, principle-learning, concept-formation, discrimination-learning, memorisation, automatization, etc.

The procedure necessary for using a system of learning tests is not as simple and readily understandable, as is teaching someone how to operate a TV set. It should be compared instead with learning to play a musical instrument or with running a computer, which have to be studied as a special art. Therefore, the effective handling of learning tests must be developed in pioneer experiments involving the intensive co-operation of both teachers and pupils. In subsequent implementation experiments it must be made teachable by routine training.

The development of learning tests is a long-term task. In interaction between instructional practice and corresponding research in educational psychology, the test materials must be continuously improved on the basis of formative evaluation, simultaneously with the further improvement of methods of application. Thus, the learning tests always need to be integrated with the other components of the curriculum, also under continual development. Included in these are the teacher's lectures, class discussions, work in small groups, games, co-operative projects, programmed instruction, etc.

In contrast to the traditional instructional materials, learning tests will not be available within the foreseeable future as "finished products". Of course this is also true of other curricular materials now under systematic development. However, one will not wait to turn them over for use until they are "completed". Rather, they will be brought out in a relatively early stage of development, as soon as they can be viewed as making a contribution to an improvement in instruction. By far the more comprehensive part of the developmental work should be carried out on the basis of continually developing evaluation in broader and more systematic co-operation between actual performance on the job and research, and in continual and complete integration with other curricular projects.

By means of learning tests, the teacher is given the opportunity to check the effects of his instruction. In this way, he is directly confronted with educational problems relevant to him personally. This continuous confrontation with actual problems encourages the teacher to come to grips in a productive manner with an action-oriented theory of instruction and with ways for improving his educational techniques.

It should be possible to offer co-operating teachers the stimulus and the opportunity of acquiring scientific knowledge and skills relevant to the respective projects in educational psychology, evaluation methodology, curriculum theory, statistics, etc. On the basis of such knowledge some teachers might acquire further qualifications for research and developmental activities.

Those teachers, too, who only utilise the system of learning tests without co-operating in its development should be given the opportunity to acquire practice-oriented scientific knowledge.

Since the pupils can immediately check the results of their learning, they are in a position to evaluate and control their own learning behaviour. Opportunities for a partnership-like co-operation and discussion between pupils and teachers arise when the methods of handling the test efficiently and evaluating its results are being studied.

It is the responsibility of the *Development Centre*, among other things:

- to develop, or to give advice on, the first draft of a learning test system and to prepare pilot projects;
- to evaluate modifications or extensions of materials and application techniques worked out in the individual learning groups, to integrate them into the system, and to make them available to all persons concerned;
- to deal with problems of organising the learning process on the basis of statistical effect analysis, theoretical reflections and special studies.

In all these tasks which, to a considerable extent, are of an evaluative kind, the centre has to co-operate both with the people engaged in practical instruction and with those doing fundamental research.

The division of labour between periphery and centre is to be more systematically developed on the basis of evaluative information gathered in the course of project work: on the one hand restrictive tutelage of teachers has to be avoided; on the other hand provisions must be made for securing appropriate horizontal and vertical co-ordination as well as scientifically and technically satisfactory project work at the periphery.

Formative evaluation within the learning test project will be based on two feedback systems.

Feedback System No. 1 (FS 1) yields immediate information eg by simple display of item difficulties at various achievement levels within individual tests or across sequences of tests. More or less

systematically gathered observations, proposals and reactions of pupils, parents and teachers constitute further information to be gained by FS 1. This feedback system enables the project workers roughly to adapt their constructional techniques, principles and assumptions without delay. In many cases such approximate corrections can lead to remarkable improvements. Over and above this, Feedback System No. 1 stimulates and supports those learning processes in teachers and pupils which enable them to participate productively in action research.

Feedback System No. 2 works at a more sophisticated analytic level. It deals primarily with those problems relevant to classroom learning and teaching which reach to some extent into the field of fundamental research.

Structuring a developmental system

As is shown by the example of the learning test project conceived along the lines of systems analysis, *the functions of evaluation, theoretically controlled research work, developmental activities, implementation and teacher training are closely interrelated. Comprehensive project evaluation, as a test of:*

- project goals,
 - theoretical foundation and technical suitability of planning,
 - practical realisation in instruction,
 - and intentional and unintentional effects,
- provides for a continuous and minimally-biased outlook on the whole problem situation and contributes, in doing so, to the integration of the various fields of work participating in the project.*

The more one succeeds in integrating the conceived goals of both experienced teachers and researchers, the more productively a co-operative developmental system will function. This should be achieved insofar as the researchers work together with teachers on the job, using formative evaluation, on the problems arising out of daily learning and teaching processes, and they then relate their research goals to the outcomes of these evaluations. And the teachers should develop their interests in educational reform in communication with researchers in a continual confrontation with reality on the basis of as comprehensive an evaluation of their classroom instruction as possible. In this way, they could emancipate themselves from hampering tradition as well as from one-sided trendy ideas.

One might expect — without being too utopian — that a more efficient educational research and

developmental system could be worked out. The speed with which this all could happen would more or less depend upon how favourable were the surrounding circumstances.

Such a developmental system needs to be based upon European or international co-operation because of its complexity and the enormous extent of the problem. The working out of means and methods for co-ordinating projects on a European basis has already been tackled and shows great promise. The Council for Cultural Co-operation has made a substantial contribution towards the establishment of a starting point for co-operative work on projects through their European trend reports, the planned contact projects workshops, and the information system EUDISED.

These beginnings call for supplementation by means of a scientifically founded strategy for the optimal choice of project goals. Such a strategy would need to be based on a catalogue of the identifiable current innovation problems in education in the various countries. Naturally, the planning of the projects will often be decided not according to scientific principles, but on a political level. Nevertheless, the political decisions can be adjusted and influenced by scientific criteria of selection worked out on a reliable basis.

The mode of operation and the horizon of the aspired developmental systems would have to differ essentially from the traditional manner of educational research and developmental activities. Until now, one dealt very often with relatively unco-ordinated and short-sighted individual projects which were "terminated" with a more or less comprehensive scientific publication or with the publication of "finished" learning materials. The perspective of such projects is usually rather narrow, that is, various components of a system are placed in the foreground, and other no less relevant components are not given sufficient consideration. The great complexity and multitude of aspects in educational problems, which are frequently also burdened by ideological considerations, appear to be difficult to master without special systems analysis techniques.

A modern research and developmental system should begin with relatively few very carefully chosen key projects which can bring about a system of complementary project associations. The way to such project associations could be opened up in an especially effective manner through the consistent application of the principle of comprehensive evaluation.

New problems appear repeatedly as a result of such evaluations. These can generate subsidiary projects within the areas of implementation, basic research,

methodology and goal determination. To bring about the necessary contacts with eligible project workers in the respective areas, a well functioning communication system and appropriate organisational provisions must be established.

Several aspects of the establishment of an integrated developmental system can be illustrated by means of the example of the two Austrian projects, already mentioned.

Selection criteria for key projects

An analysis of the innovative relevance of these projects yields two main types of criteria for the selection of key projects:

First, criteria which relate to the suitability of a projects for the mobilisation of innovative interest and for initiating learning processes within the important groups of persons involved in educational innovation. One would need to examine in particular how one could best awaken the interest of parents, teachers, scientists, administrators, politicians and the general public, and which learning goals are primarily to be attained by these various groups.

The example of the Austrian comprehensive school experiments leads to the assumption that projects which change the organisational structure — in contrast to attempts at changes in teaching methods for instance — arouse especially great and very widespread interest right from the beginning and also gradually lead to a consideration of a variety of educational problems.

A second type of selection criterion for key projects relates to the suitability of a project as a core for the establishment of a project association.

Core projects presumably will have to have goals with far-reaching significance and be directly applicable to the school setting. They should be able to be organised according to principles of action research and require permanent co-operation between practicing teachers and researchers of various disciplines.

A comprehensive formative evaluation to examine all components of the project should continually lead to the identification of essential problems which then can be dealt with in subsidiary projects.

Core projects are expected to lie mostly in the field of curriculum development. Here teachers and researchers can productively meet one another in the realistic setting of classroom learning and teaching, whence the problems branch out.

The evolution of a project association

The Austrian learning test project can be viewed as an attempt at the establishment of a core project it should consist of co-ordinated partial projects carried out by separate project teams who develop learning test systems as components of curricula for various school subjects. Developmental work has already begun for mathematics and English as a foreign language.

Taking the projects for mathematical learning tests as an example, I can only allude to the way in which subsidiary projects are beginning to become linked to the core project. These were to be worked on by specialised project teams in communication with each other and with the core team.

Pupils with various aptitudes demonstrate great differences in speed of learning. These differences may vary in size with respect to various types of learning goals: e.g. low achieving pupils appear to be able to achieve goals in mathematics at the lowest level of the Bloom Taxonomy relatively easily. At the same time, they require a great deal of time for certain other types of mathematical goals which generally lie at higher levels. The differences in learning ability give rise to various problems in the learning of mathematics. Of these, only two will be pointed out here:

The first problem concerns the sequencing of the learning steps for pupils at various levels. In this case, the following questions arise:

- Are there different optimal learning paths for reaching certain learning goals in mathematics for pupils of varied aptitude levels? Or does the proportion of their learning time remain constant even when the learning path is varied?
- If there are optimal learning paths for specific aptitudes, which characteristics of the learning path or of the individual learning steps can serve as criteria for the optimisation of the learning path for pupils of various aptitude levels?

The second problem area connected with the differentiated levels of the pupils relates to the question: how should the final goals for mathematical instruction be reconciled with the various aptitude levels? For example, when the learning time differences between the levels for various types of learning goals surpass a certain standard, then the acute question arises: how much learning time should be spent on the various learning goals at the individual aptitude levels, in order that the optimum set of goals can be achieved at each level?

Further problems arising out of the prospective core project "Learning tests for mathematics" can only be briefly mentioned here:

- The development of techniques for the analysis of the learning process based on learning test results. Above all, at issue is the working out of criteria for the optimal sequencing of the learning steps.
- An analysis of the relationship between the logical and the psychological structure of mathematics with special emphasis on a critical examination of the significance of the so-called "New Math" for the learning of mathematics.
- Working out criteria in order to determine final aims for mathematical instruction.
- Integration of the learning tests with the remaining components of the curriculum, particularly with the co-operative and autonomous forms of learning.
- Integration of the learning of mathematics with learning in other areas.

This list of research and developmental problems reveals only a small selection of the multitude of problems which have already become apparent. Scarcely one of them has a practically applicable scientific solution. For that reason these problems had to be treated in corresponding subsidiary projects.

It does not follow however, that the core project has to wait for its developmental work until the subsidiary projects or other research have brought about usable results. On the contrary, the core project produces results on the basis of the best information available at that particular time and uses the results of subsidiary projects and other research as soon as they are available.

Summary

This attempt at sketching an integrated developmental system intends no more than to give an initial estimate of a course to be taken. The techniques for the planning and organisation of such a developmental system require comprehensive evaluation in order to bring about continual improvement and to accommodate to current scientific and social developments.

We will need to adjust ourselves to the fact that spectacular results can seldom be achieved in short spurts in the area of education. Gradual but sure progress can be achieved on a broad front if one makes far-sighted plans, and then, on the basis of

patience and adequate resources, builds up a co-operative developmental system controlled by comprehensive evaluation.

The educational system now under development can be viewed as a highly complex cybernetic system with many subsystems to be regulated. It is not only a question of continuous corrections of deviations of actual values of the system function from desired values by resources inherent in the

system, but also of controlled development of the components of the systems themselves: the sensors, the feedback mechanism, the executing body, and the method of analysis for the desired value.

It is the job of a comprehensive evaluation as a systems component of developmental projects to continually insure the information basis for such an extensive, co-ordinated, controlled development.

Summing-up lecture

In all the countries whose experience we are discussing here, the scientific evaluation of pilot projects for school reform is in the initial experimental stage. This has been emphasised and illustrated by the lectures and participants at various points. In their statements on this experimental stage the speakers have referred to the devising and testing of appropriate methods and of the instruments whereby they are applied as well as to the establishment of organisational models on which to base the evaluative research projects.

We have only to bear this in mind to realise that the symposium could not be expected to furnish ready-made solutions, much less practical directives.

It seems to me, on the other hand, that we were justified in expecting that on the basis of their accounts of the theoretical research or practical experiments carried out by them the lecturers would develop ideas which would be of use for the further development of theory and the practical evaluation of pilot projects for school reform. The justification for this is to be found in the discussion groups, each of which took up in its own way the ideas put forward by the lecturers and developed them further by integration and contrast in equal degrees.

It is not surprising that all group discussions diverged repeatedly from the main theme and extended to general problems of educational innovation and decision-making processes, seeing the interrelations between these questions and evaluative research.

These interrelations must also be apparent from the group of basic questions on which I should again

by Professor W. MITTER,
Deutsches Institut für Internationale Pädagogische
Forschung, Frankfurt/Main

like to dwell now, in winding up. It is not my intention to assess the meeting but to summarise my own direct impressions as a participant and a learner.

I will be very brief and confine myself to three groups of questions which seem to me to create dialectical tensions — dialectical, in that the phenomena involved may be considered as antitheses which pervade educational, scientific and social processes, neutralising each other and thus giving rise to new tensions.

I refer in the first place to the conflict between centralised and decentralised innovation in school reform. Most of the lecturers came out in favour of school-based, decentralised educational development and emphasised its functional advantages as well as its emancipatory ones to which I shall revert. Not only Mr. Voigt, as a representative of school administration in a position of political responsibility but the other speakers too dwelt on the need for central overall direction of school innovation, though sometimes in a modified form. This led — I believe rightly — to doubts being expressed about the existence of a general trend towards decentralised innovation, and the statement on the subject by the Minister of Education of Land Hesse might be regarded as virtually the "official evidence".

It would surely be an over-simplification, and tantamount to devitalisation, to regard direction of central national or international, as it will no doubt increasingly become — education policy by state or public bodies merely as a necessary evil or a disturbing factor. That would mean radically overestimating the capacity of decentralised model

schools to prepare pupils for a share in solving general social problems.

The second conflict is between the two different kinds of research repeatedly referred to here: action research as an ally of decentralised school development and traditional empirical research, whose continued importance for the comprehension of macro-structures and the compiling of data for the reform of education systems has not been questioned. Mr. Klafki summed up in his report the findings of his group, which indicate the contribution made by action research to defining and differentiating fundamental scientific research in education and emphasise the importance of traditional empirical methods for action research. The significance for action research of the fact that only the efforts to overcome a false polarisation of attitudes have cleared the way for further development of both types of research has become obvious from the example of the controversy over the isolation of variables.

I now come to the third group of questions which includes the interpretation of the fundamental concept of "democracy", its reflection in education and the school and participation in decision-making processes by teachers pupils, parents and other persons closely concerned. This is what is meant by the term "Demokratisierung" in German; our neighbours content themselves with less ambitious and less overall definitious (such as "participation" in French).

The discussion on these questions has revealed perhaps most clearly the great differences in fundamental attitudes and their relation to the historical traditions of the individual national states with their specific implications for education.

From the relationship between researchers and teachers engaged in the scientific evaluation of school experiments it has become clear that the two kinds of activity are in fact convergent and above all of equal value. This does not mean that the field should be left to the utopian ideas of all-round experts; in practice this might quickly lead to a preponderance of all-round dilettantes.

With reference to my thesis of the dialectic in the three conflicts outlined here I should like to add a personal comment. In Eastern Europe the concept of "practitioners taking part in research" has been introduced into discussions on educational and social forecasting: these practitioners include teachers. Without going into the question of comparability, I consider that this concept comprises factors which could also become pertinent in our social systems. In this connection we should remember

the need, already referred to, for teachers, parents and pupils to become conversant with the fundamental principles of research methodology and its instruments. However, the question put at the end of the plenary session remains unanswered: to what extent can the scientific nature of the work be safeguarded within the context of such a research-structure?

This concludes my survey of the three groups of problems. Allow me to make one concluding comment which is connected directly with the question of "democratisation". It was natural that in this symposium, which was centred on the open and controversial questions of scientific evaluation of pilot projects in school reform, there should be intensive analysis and discussion of the crucial role of the teacher. The pupil's participation in the research process, an admittedly highly complex subject of investigation and experimentation related to general and particular factors such as physical constitution, age group, social origin and the like, was necessarily left in the back-ground if not overlooked.

In making this comment at the end of my talk my aim is to give it its due weight. For I have learned from my own experience that both behaviourists and educationalists who derive their ideas from radical criticism of our society make all kinds of ingenious statements at their meetings about systems, methods and prospects, however much their individual philosophical and political standpoints may differ. On the other hand, the word "child" is often dropped in passing, at best, if at all. I should like to relate this comment to two questions raised during the symposium.

First there is the question of teacher participation in action research. Is it altogether impossible that teachers, in generalising their own ideas, would lose touch with the actual learning and life situations of their pupils and demand too much of them? We must see this danger and must prevent it.

The other point is the need, mentioned by Mr. Gauthier in discussion, for common cores in decentralised and individualised education. They are certainly necessary, not only to protect the pupil against unreasonable strain when changing schools but also to enable him to acquire the necessary general social knowledge. But over and above this content aspect there is the question of promoting learning processes which — simultaneously with and in addition to the acquisition of generally necessary social knowledge and knowledge related to the individual — take account of affective factors as essential to individual self-determination.

REPORTS OF THE DISCUSSION GROUPS

Group discussion I and II

Rapporteur: Mrs. B. GAARDER,
Research Consultant,
The Royal Ministry of Church and Education, Oslo.

Evaluation of school reform projects

The group experienced great difficulty in responding positively to the first theme. Following a well-known procedure in education, the group therefore proceeded from the particular to the general and found wide variations in the countries represented. Some examples follow.

There were projects which might be described as local pilot studies of a kind which, if replicated in other localities and carefully evaluated, could influence central policy regarding school organisation. There was an example from one country of a great deal of money being spent on decentralised projects with results of so limited a nature that the central government was currently playing a more active, though not a prescriptive rôle.

In other cases where national projects were tried out in many different localities, great attention was paid to the multiplicity of factors which can influence outcomes.

From discussions of these and other examples the group reached the conclusions which follow.

— The dichotomy between centralised and decentralised approaches was not as clear-cut as the title of the symposium implied. It might well be considered necessary in some countries to differentiate between those research and development activities which were appropriate to (a) centralised, (b) decentralised procedures. The group found itself unable and unwilling to produce guidelines which would be universally applicable.

— Educational policy in all countries appears to result from complex interactions between a number of political, economic and social forces, with the findings of educational research commonly playing a relatively minor rôle. However, the group felt that although governments would not necessarily postpone decisions pending the results of long-term research, it would be valuable for them to initiate short-term feasibility studies with the aim of providing reliable information on which to base future decisions. The group recognises that, on

major issues, educational research is but one component in a totality of factors influencing the decision-making process.

— Policy decisions in the field of education are customarily formulated in somewhat broad terms. It may be decided, for example, that provision for nursery education should be extended to all children from the age of three. Once a decision of this kind has been implemented, an important function of research and development work is to build up a many-faceted picture of the conditions and circumstances most apt to maximise the potential benefits of the reform, so that any subsequent modification of the original decision may be soundly based on empirical evidence.

For example, if it had been decided that nursery education should be extended to all three-year-olds, a programme of research and development would be necessary in order to provide information on questions of detail such as: Should nursery education take place in separate schools or in classes/units attached to existing primary schools? Does nursery education exert a differential effect on children of differing socio-economic backgrounds? Can physically-handicapped children be successfully integrated into nursery schools/classes? And so on. Well-defined, systematic and properly evaluated research and development studies are essential if broadly-conceived decisions are to be given necessary depth and texture.

— An evaluator who is a member of the development team seems to be particularly important at the start of the project. His task then is to collect data for a quick feedback of information about the progress of the project. At this stage techniques of data-collection may lack the refinement normally considered necessary in rigorous research. These data might be qualitative or quantitative. The information fed back to those taking part in the study needs to be put in an easily understandable form.

As the project develops further external evaluation is also needed. The perspective in this case would be wider. External evaluators can underline important factors of the innovation which have been

overlooked or taken for granted. External evaluators might also be better able to summarize the results and see more possibilities for generalisation.

— A major area of agreement within the group was the acceptance of the view that both in developmental and evaluation situations the involvement by participants should be seen as being important to the overall outcome of any specific programme. An analysis of the apparent failure of many development programmes in the past has resulted in increasing efforts to involve more teachers in the creative processes called for in such work. This experience would suggest that in corresponding fashion it is important to involve as many teachers as possible in the evaluation process. Thus a major aim of the educational research programme should be to promote a greater awareness on the part of teachers, both critical and creative, of educational issues at every level. This in

turn should help serve to raise the whole level of public debate on educational matters so that important decisions regarding the rôle of education in society may come to be made on the basis of fuller information and involving a more critical assessment of the issues involved.

The group agreed that irrespective of the style of evaluation adopted an important function of evaluation should lie in in-service training whereby the information gathered would be critically examined with teachers in relation to the programme under consideration. A further important rôle of the evaluator must be that of an impartial but critical commentator whose views can illumine the information available to the decision-makers. These may be teachers involved with detailed curriculum matters, parents concerned with the educational progress of their children, or politicians acting in issues of major social importance.

Group discussion III

Rapporteur: Mr. B. W. KAY,
H.M. Staff Inspector,
Department of Education and Science, London.

Relationship between teacher and researcher

Teachers, once enthusiastic for reform, are now sometimes reacting against innovation and research, especially when it is imposed from without and when it has no obvious benefits to confer upon their schools and their pupils. Consequently it is sometimes difficult to obtain the co-operation of teachers, and in particular of their Unions, and this is particularly evident when the research involves questionnaires or externally-imposed tests, which can be interpreted by teachers as 'spying' on them. The researcher is thought of as an 'outsider' even if he may have only lately left the classroom himself. In one country the hostility by the Ministry to some forms of research discourages teachers from participating.

The researcher can however improve the attitudes of teachers in a number of ways.

- In short-term applied research projects, he can make explicit the benefit to pupils — this is rightly a top priority for teachers.
- He can take pains by involving them in planning etc. that teachers and pupils do not feel that they are being used as guinea-pigs.

- He can ensure full preparation and discussion, involving teachers and others, before research is begun. Some countries have set up consultative machinery involving teachers and their unions.
- He can use volunteers as far as possible.
- Some countries provide extra resources for the schools involved and extra pay for the teachers. Participation in projects can also bring prestige.
- In-service training can help to give teachers confidence in tackling the unknown.
- Research is most acceptable to teachers when it most obviously is designed to meet the felt needs of school or society.

Experimental schools

Some innovation can take place in some areas of the school without greatly affecting the school as a whole: some involves a change in the whole structure. Some schools become 'de facto' 'experimental schools' by the enthusiasm of their staff over the years for taking on new projects. However the term is best kept for those schools involved in a

planned and properly evaluated research exercise. The problems of comparison with 'control' schools are severe where experimental schools attract extra resources, greater prestige and so better teachers and more interested parents. In some cases these problems are met by similarly benefiting the control schools.

Motivation for research

Teacher-felt needs are one valid basis for research, particularly in the field of curriculum development. But this can result in piece-meal work, if there is no long-term coordination and planning.

Another basis is society's needs. In a time of change it is difficult to anticipate the needs of ten years hence — though this is the sort of time-scale that research must follow. But the extent and rate of change can both be over-stated — more remains stable than suffers change.

Education serves to

- transmit an inherited system of values — this remains an important and stable element;
- teach the skills needed for a person to live adequately in a sophisticated society; these include a range of subskills which may change in relative importance;
- convey the non-cognitive attitudes and qualities needed for a mature, full and happy life in society;
- prepare pupils for change by encouraging flexibility of approach and by preparing them for life-long education;
- anticipate as far as possible future vocational needs — some pupils at present in school will be in jobs which do not exist at present. In some countries change is taking place too quickly for the results of research to be waited for.
- Society in some countries is seeking reassurance that educational institutions are achieving adequate standards for their pupils. Hence a new demand for evaluation.

Decision making in the setting up of research

Politicians on the one hand and teachers and others concerned with education on the other are the agents mostly involved in interpreting the educational needs of society, and where they conflict, in deciding priorities. Neither necessarily fully reflects the views of a pluralist society. It is desirable that both should have contact with representatives of varying interests and be responsive to them.

Apart from the basic educational requirements acknowledged by all the identification of needs is to some extent subjective — the interpretation of society's needs is basically political, from whatever source it comes.

The politician is the main patron of research, but his motives are not always pure. Research can be called upon to support a decision already made on political grounds rather than to give guidance in coming to a decision. By financing research and awarding priorities the politician is in effect manipulating research in his own interests. The researcher may moreover be called in to test an hypothesis formulated by someone else, and possibly expressed in a form that he would not have chosen. A moral dilemma can arise over the publication of material whose conclusions run counter to government policy. In some countries the researcher has the right to publish the facts under any circumstances, but not his interpretation of the facts.

It is vital that autonomous sources of research like institutes and universities should continue to exist. Even this cannot however be regarded as being free from political bias, and in fact they may lack the direct responsibility that the politician owes to parliament and electorate. Even here the issue of publication may arise.

Dissemination and its implications for evaluation

There appears to be an inevitable loss of efficiency as projects become more widely disseminated, and this loss of efficiency begins as soon as the project passes beyond a few closely involved schools. Not only does teachers' motivation become less, but teachers not closely involved in the setting up of a project often fail to understand fully the aims and methods involved.

This loss can be minimised by the setting up by the project team of an in-service training programme for teachers starting to use the material, by the provision of very full printed guidance for teachers in the form of a manual, by the encouragement of the setting up of local groups of teachers working closely together.

Motivation can also be maintained after the formal end of a project where teachers — and even pupils — are involved in continuing assessment. This can also provide them with a continuous feed-back.

It is difficult to ensure the 'multiplication' effect since teachers often fail to communicate their enthusiasm and expertise to others. However projects have to expand with limited manpower and the best way of maintaining the impetus appears to be

the forming of local groups, which can maintain closer contact between researcher and teacher. Guidelines from the centre may still be necessary.

The nature of evaluation

The range of evaluation approaches includes at the one end objective scientific description (which will however often imply value judgments, as the evaluator brings his philosophy with him). At the other end it will include standardised testing of easily measured data with a high degree of accuracy. A range of instruments will be needed to cover all the proper objects of investigation, and different standards of accuracy will be appropriate for different parts of the spectrum. The more comprehensive the project or other object of enquiry the more global will have to be the approach to evaluation. The evaluator will aim at the greatest accuracy and greatest objectivity compatible with the object of evaluation. Examples were given of a range of different methods of evaluation being applied to the same situation — e.g. a series of accurate descriptions of observed processes could help to explain the reasons for different levels of achievement recorded in tests.

An example was given of a feasibility study shortly to be mounted to evaluate primary education by a sampling technique using three lines of assessment.

- standardised tests;
- structured subjective judgments by inspectors, using four or five point scales in a questionnaire prepared jointly by inspectors and a research institute;
- descriptions of educational processes observed by inspectors.

In order to standardise the second and third as far as possible, "moderation" would be carried out by including in the team visiting each school a member of the team which had originally drawn up the criteria, or one who had worked with such a member. The purpose was to evaluate a broad spectrum of educational objectives.

In action research there is room for both empirical quantitative evaluation and for less precise measures. The researcher needs to be aware of the strength and limitations of his various tools.

- He should make use as far as possible of quantitative techniques, but there are some areas of educational objectives which such techniques cannot easily encompass. It may then be better to use less reliable methods rather than to ex-

pend disproportionate efforts to obtain only meagre additional results by quantitative methods.

- The involvement of teachers as collaborators rather than consumers brings advantages to counterbalance some loss of scientific precision. (Participant observation, incidentally, is not a recently introduced technique). There are areas in which the teacher actively involved can evaluate objectively, but an overall judgment is not compatible with close involvement. There is a danger that an ideological bias may be imported.
- An insistence on precise measurement can limit the range of evaluation. Even instruments like questionnaires are fallible because of the varying attitudes and degrees of cooperation of the answerer.
- Any evaluation influences what is being evaluated — even observation affects the situation observed and could falsify the data.
- What the observer sees can be limited by his preconceptions. The evaluator might well adopt in the first place a very open attitude and only later attempt to structure his observations in the light of experience.
- Observation can result in more or less quantified information a three or five point scale is a crude measurement; a statement that "so many examples of this or that were observed" may be more precise.
- Although some qualities to be evaluated are very difficult to define and may be impossible to measure, some definition, however broad, is needed before evaluation can be attempted.

An example was given of an attempt to evaluate pupils' behaviour in a number of areas selected from teachers' identification of their educational objectives. Qualities such as perseverance or critical sense are broken down into a series of specific, easily identifiable items of behaviour. The intention is that a fellow-teacher should observe a series of classes and note down entries on an observation schedule. It is inevitable that teachers will consciously or unconsciously adapt their methods to produce examples of the behaviour under observation, and that the pupils too will be influenced. This is recognised and accepted. This is regarded as a piece of formative evaluation which will tend to modify and improve teaching techniques.

In another country a project for the evaluation of five pilot schools in a number of fields of enquiry failed because of inadequate initial planning and

subsequent changes of policy. In place of the global project the research team is now restricting the evaluation to a more limited field. In particular groups of volunteer teachers, helped by psychologists, are studying four areas of the curriculum and defining objectives in terms of the behavioural qualities which it is hoped pupils will acquire through the study of subjects, rather than in terms of the information content of subjects.

In the course of discussion a range of views was expressed about the validity of inter-system comparative evaluation. The results of such evaluation are particularly likely to be misinterpreted or misused. In practice only a few characteristics shared by the two systems to be compared can be selected, and these are frequently selected more on political grounds than because they offer the best basis for overall comparison.

Costs of evaluation

Evaluation involves expenditure in two ways:

— The actual cost of the evaluation exercise must be proportional to its product — some expensive

evaluation procedures have barely justified their cost — the same conclusions have been reached independently by more simple and direct methods. It is important that *unnecessary* duplication of research should be avoided — hence the importance of ERIC and EURISED.

— New methods of teaching and new forms of school organisation arising from research can be more costly than traditional ones in building requirements, teaching resources, in-service training and supply of teachers. An element of costing for the implementation of a project should be built into the evaluation so as to give better guidance for policy decisions by Ministers.

Terminology

It was recommended that for future conferences of this sort among the preliminary papers should be a glossary in the working languages of terms likely to be used. Where no verbal equivalent exists, a short description or definition might be required.

Group discussion IV

Rapporteur: le Recteur M. GAUTHIER,
Inspecteur d'Académie,
Ministère de l'Education, Paris.

The multilingual group was composed of representatives of the following States: Austria, Belgium, Federal Republic of Germany, France, Italy, Norway, Spain and Switzerland.

The problems it studied may be classified according to four main themes:

- decentralised innovation,
- action research,
- formative and summative evaluation,
- evaluation techniques.

Decentralised innovation

Innovation may be initiated either by the central government or at the local or regional level. In the case of an initiative by the central government deciding to survey a reform project, either the same research can be entrusted to a certain num-

ber of school establishments, or different elements of the research can be distributed among them. Several participants stressed the importance of local initiative in innovation. They consider that the main function of the central authorities is to listen to teachers, who must play an essential part in defining the objectives of innovation.

The relation between the central authority and the units on the periphery was examined with regard to the nature and scope of the control exercised by the central authority, and with regard to the degree of freedom to innovate accorded to the local educational units, which varies according to the question under consideration.

In the matter of *school structures*, reform projects almost always derive from the central government, because of the need to assure an indispensable coherence in the educational organization of the country, and because this brings into question the State's general and financial policy.

As for *teaching content*, in some countries it is imposed in systematic fashion, while in others schools and sometimes even the classes themselves are allowed to exercise choice in relation to a more or less important portion of the curricula. But this latitude is always limited by the need to prepare pupils to pass examinations in the subject, and by the concern not to place at a disadvantage the growing number of children who have to change schools because of family mobility. It is therefore necessary to impose at the very least a common core curriculum; on the other hand, some countries are turning to the concept of more flexible examinations.

It is in the field of *teaching methods* that one finds the greatest scope for initiative, although it should be noted that innovation in teaching methods can lead to a modification of content (for instance language-learning by audio-oral method brings about a new type of knowledge and practice of the language).

Central authority action is felt in several ways:

- examination of local initiatives and the necessary consent for setting up the experiments,
- possible control to prevent local initiatives contrary to the country's educational policy from slowing down instead of promoting the evolution of the educational system,
- coordination of decentralised research,
- help for local units engaged in research,
- setting up of a reciprocal flow of information between the central level and the establishments at the periphery (exchange of information about experiments and results, ideas bank),
- financing, either by the State alone, or by division of the expenses between the State and local or municipal budgets,
- final evaluation and generalisation decisions.

Action research

The group aimed to define it in contrast to the more classic form of development research, a procedure consisting essentially in research limited to a restricted group, at the end of which the problem of possible generalisation arises.

The nature of action research is wider and more dynamic. A greater number of teams in the field participate in the study of an innovation project. These teams are made up of working teachers helped by specialists in the relevant disciplines and by research workers, who together establish a process of direct interadjustment of practical experience and theoretical work.

Can one say that action research must be democratic, and is this qualification compatible with the scientific nature of the research? I would seem to be a question of involving the greatest possible number of people affected by the reform project, of obtaining the full participation of teachers, pupils, parents and even locally elected officials, in fixing objectives, in elaborating methods and curricula, and in observing the conduct of experiments.

Furthermore action research has a very strong internal dynamic, for decentralised experimentation carried out in a large number of local units requires periodic comparisons by the different teams, who go on by means of successive syntheses to the progressive elaboration of objectives and methods. It is a question here of a continuous process of internal evaluation within the teams themselves, who are thus led to adjust their actions to the objectives, or to correct and modify these objectives in the light of difficulties encountered.

Formative and summative evaluation

Besides the continuous formative evaluation organically involved in action research, we must distinguish summative evaluation, which interests the central authority responsible for making decisions concerning the generalisation of the tested reform. By whom and how can it be carried out? Should it be carried out by external evaluators? The group considers that summative evaluation should not be divorced from continuous evaluation, and that if external evaluators are brought in, it should proceed on the basis of the work of teachers and researchers who have taken part in the experiments. It should invite their participation or risk being challenged.

It was noted that in action research decentralised cover in a large number of "fields", this multiplicity of test locations is both a source of wealth owing to the number of contributions and a source of difficulties for the evaluation because of the diversity of conditions to be found in the various experimentation locations. Rather than provide a value judgment on the tested innovation as a whole, summative evaluation would consist in giving an account of the various forms of research and of the results obtained in the various schools where it was carried out, bearing in mind the conditions and mode of work peculiar to each.

The group discussed the recruitment and training of research workers. Besides academics trained in institutes of educational science, recourse is most often had to psychologists, sociologists and teachers who devote themselves temporarily to the task.

Since they have seldom had previous training, these researchers acquire this training empirically during the research process, which increases both the duration and the cost of the experiment.

Evaluation techniques

Following the general discussion on evaluation techniques, the last meeting allowed the group to examine concrete methods of evaluation applied in certain school experiments in progress.

The impression gained from the various contributions was that preoccupations were similar and

that similar means had been employed. Among the procedures used and remarked on were the elaboration of pupil maturity tests, sociometric inquiries, the partenal situation, and performance tests taken at the beginning and during the course of study.

In addition, there was agreement in acknowledging the need to inform everybody affected by the innovation as completely and as coherently as possible — teachers, pupils, parents — as to the objectives of the innovation being tried out.

BACKGROUND DOCUMENTS

Austria

PART I

THE EVALUATION OF COMPREHENSIVE SCHOOL EXPERIMENTS

Within the framework of the programme for Austrian school development, pilot projects are being carried out in the schools for 10 to 14 year olds, according to a federal law enacted in 1971. The purpose of these projects is to test forms of comprehensive schools.

At the centre of the experimental work stands a type of integrated comprehensive school which provides for attainment groups (setting) in mathematics, German and English from the 5th to the 8th grades. Heterogeneously grouped classes are used for instruction in all other subjects. In addition to this, a type of school devoted to "guidance" is being tested. This is conducted like an integrated comprehensive school in the 5th and 6th grades only.

The law prescribes scientific control of the comprehensive school experiments so that information can be gathered which is to serve as a basis for political decisions concerning the organisational set-up of the schools for 10 to 14 year olds. For the time being, the experimental types are to be tested for five subsequent years from the 5th grade up to the 8th grade respectively.

by Professor G. PETRI,
Zentrum für Schulversuche und Schulentwicklung,
Graz.

Evaluation goals

First of all, information concerning the further development of school types and evaluation systems is being worked out (formative evaluation).

In the final phase of the experiment, data are to be gathered on which to base a decision about the organisational form of the schools for 10 to 14 year olds. For this purpose, hypotheses about the differing effects of the traditional school as opposed to the experimental school types are to be tested (summative evaluation).

For the preparation of a decision, essentially the following questions will have to be examined:

- Does the new type of school organisation provide a better education for pupils with a socio-economically and socio-culturally disadvantaged background?
- Do educational opportunities improve in those regions with an insufficient number of higher level schools?
- Do pupils attain, on the average, a higher standard of education when strengths and weaknesses are taken into consideration by a setting systems?
- Considering the fact that pupils of all attainment levels are together for a greater part of the instruction time, is it possible to develop

more satisfactory contacts and a better understanding among pupils with varied social backgrounds?

- Is the repeater problem satisfactorily solved by the new form of school organisation in which repetition of a class is avoided by reclassifying the pupil to the next lower attainment level in the respective subject?
- Does the new type of school organisation, which substitutes furtherance of the individual for the principle of selection, promote socially-integrating behaviour patterns on the part of teachers and pupils? Does it increase the motivation of the pupil to learn more about the subjects?
- Does the new system, as opposed to the traditional one, lead to a lowering of attainment on the part of pupils with certain levels of aptitudes?

Methods

The evaluation of the school experiments proceeds from detailed hypotheses about directly observable facts. Each of these hypotheses makes comparative statements concerning:

- differing effects of
- various school types on
- pupils with specific characteristics
- under specific educational conditions.

The following itemisation offers a rough overview of the variables which come into consideration in the building of hypotheses:

1. Dependent variables:

- Career
 - + Educational level attained at the end of compulsory school attendance
 - + School or occupational training after the 9th grade
 - + Vocational or scholastic level attained later on
- Cognitive variables:
 - + Attainment in the subjects with attainment levels (mathematics, German, English)
 - + Attainment in the subjects taught in heterogeneous groups
- Non-cognitive variables
 - + Motivation of the pupils with regard to learning and education
 - + Attitudes towards school
 - + Stress on the pupil, fears and problems in connection with school

- + Effects of social learning
 - Sociometric structure of pupils' groups
 - Forms of interaction between pupils of varying backgrounds and capabilities
- + Attitudes of the parents towards school
- + Teacher attitudes

2. Treatment variables

- Traditional schools
 - + Grammar school (Gymnasium)
 - + Secondary modern school (Hauptschule), first stream
 - + Secondary modern school (Hauptschule), second stream
- "Integrated" comprehensive school
 - + Attainment level, group I
 - + Attainment level, group II
 - + Attainment level, group III
 - + Instruction in heterogeneous groups
- Guidance period
 - ...
- "Additive" or "multilateral" comprehensive school
 - ...

3. Control variables

- Pupil characteristics
 - + Sex
 - + Psychological characteristics (intelligence, motivation, etc.)
 - + Social status
 - + Out-of-school help with studies
 - + Attainment at elementary school, degree of maturity for secondary school, class repetition
 - + Milieu: town, rural area
- Instruction factors
 - + Learning objectives
 - + Instructional methods
 - Textbook
 - + Distribution of attainment in the class
 - Mean
 - Standard deviation
 - + Group dynamics of the learning group
 - + Teacher characteristics
 - Training
 - Sex

- Age
- Educational style
- Teaching skill
- Attitudes

The testing of hypotheses presupposes:

- valid and reliable methods for measuring the effects;
- the uniform implementation of the various steps whose effects are to be determined;
- a sufficiently exact control of all variables which influence the effects examined apart from the experimental variables.

Of these prerequisites an adequate treatment of the control variables is the most difficult to realise in the study under consideration.

The control variables either have to be kept constant, to be systematically varied, to be strictly randomised, or they have to be taken into account mathematically on the basis of measurements.

Randomisation is of essential importance particularly for the elimination of the extremely heavy influence of teacher personality and the dynamics of the learning group.

The "Hawthorne effect" can presumably be reduced by extending the school experiments over a longer period of time and increasing their number. Systematic interviews of teachers at experimental and traditional schools allow well-founded assumptions to be made as to the distribution of attitudes which might favour a "Hawthorne effect".

Planning

The experimental phase, in which the new organisational school types are to be tried out under scientific supervision, will be preceded by a developmental stage.

In this developmental stage, the basic materials with which the teacher has to work are to be developed and tested. These include guidelines and data concerning the grouping and regrouping of the pupils into the attainment groups, as well as the mode of instruction in the various attainment levels and in the heterogeneous class periods.

The trials in the developmental stage are to be evaluated with the following goals in mind:

- obtaining feedback as a basis for planning further developmental activities;
- building up both an organisation and a staff for the scientific evaluation (control) of the experiment;

- development and/or testing of methods of data compilation, data processing and statistical analysis;
- determination of the conditions to be fulfilled to enable school experiments to be continued in a scientifically controlled manner.

At the beginning of the 5th grade, the control data are to be gathered in an entrance survey. Later, many of these data would no longer be available. Primarily, they include psychological characteristics of the pupil such as intelligence, past learning progress, attitudes, etc.

Towards the end of the 6th and 8th grades data on the dependent variables are to be gathered (scholastic attainment, attitudes towards school, learning motivation, sociometric data, etc.). The treatment and control variables which change in the course of time are also to be determined; e.g. which attainment groups the pupil belongs to, participation in tutorial sessions, learning help outside the school, etc.

The evaluation of the school experiments requires:

- a determination of the probability of the existence of varied effects of the individual school types;
- if possible, information on the size of differences in effects;
- and, if necessary, a determination of the form of the functional relationships between effects, treatment and control variables.

First results

The experiments evaluated so far belong to the introductory phase of the overall experiment envisaged, both with regard to the organisational types tested and with regard to the methods of scientific control used. The first results related to the 5th grade of the experimental schools in the school year 1971/72.

In the autumn of 1971, the Austrian school psychologists undertook a survey of 7 000 pupils in 145 experimental school classes and in 79 traditional school classes. Included in this survey were intelligence and scholastic attainment tests, a questionnaire for the pupils as well as one for their parents, and a sociometric test. In addition, the grades received by the pupils in primary school and the assessment of the pupils' maturity for the first or second stream of the secondary modern school (Hauptschule) as judged by the primary school teacher (who has the pupil all 4 years) were registered.

An intermediate investigation was carried out in June 1972 and 1973 by the school psychologists. This was performed at the end of the 5th and 6th grades, respectively, on approximately 4 600 pupils in experimental and control schools. The pupil's performance in mathematics, English and German, as well as his attitudes, interests, and sociometric variables were investigated.

First of all, the effects of school organisation on school performance and on the equality of educational opportunities were analysed (1).

The following hypothesis was tested:

Scholastic attainment of the pupils in the 5th grade in the ability-differentiated subjects (mathematics, English, and German) is, on the average, greater in the integrated comprehensive school than the attainment of comparable pupils in the traditional school system,

since:

pupils with equal aptitudes learn approximately the same amount in the appropriate attainment groups of the comprehensive school as in the traditional school (sub-hypothesis 1);

however, pupils in the comprehensive school are better allocated to attainment groups coinciding with their respective aptitudes than in the traditional school system (sub-hypothesis 2);

equally talented pupils learn more in homogeneous learning groups corresponding to their respective levels of attainment than in classes less well suited to their respective levels of attainment (sub-hypothesis 3).

Sub-hypothesis 1 was confirmed by the experimental results gathered to date: the upper level groups of the comprehensive school attained equally high results in mathematics, English and German as would be expected from equally talented classes of the Gymnasium. In evaluating these results, one has to consider the fact that various factors which could influence pupil performance in favour of the experimental school or in favour of the Gymnasium were not controlled in the present study; or, if so, only insufficiently. Such factors include: promotional measures to assist the pupil either at school or at home, the family, learning motivation, sequence in which the subject matter was taught, method of instruction, effectiveness of the teacher, number of pupils in the various classes, Hawthorne effect, etc.

It is not possible, at the present time, to form a scientifically sufficiently well-founded judgment on whether the uncontrolled factors, as a whole,

biased the results in favour of the experimental schools, or in favour of the traditional schools.

Sub-hypothesis 2 was also confirmed. In fact, it was determined that there are often considerable intra-individual ability differences in the three subjects examined. Approximately half of the pupils could not be put into the same groups in mathematics, English and German, with three attainment groups for each subject.

It was observed that there is considerable changing up and down between the attainment groups in the 5th grade. Approximately 26 per cent of the pupils changed sets in mathematics, English, or German from beginning to end of the school year.

The réclassification carried out in the 5th grade, however, did not seem to have essentially improved the allocation of the pupils to attainment groups. Those pupils transferred to higher groups showed no more marked improvement in performance at the end of the school year than did those pupils transferred from higher attainment groups to the lower levels.

This is considered as a demonstration of a need to base future reclassification procedures on psychological tests. These are already being developed. The new technique is being evaluated at present.

A clear statement as to whether a more exact placement of the pupils in attainment groups will improve their scholastic performance (sub-hypothesis 3), cannot be made on the basis of the data of the study. It is possible that the influence exerted by the degree of homogeneity of the learning group on individual performance will become more apparent in later grades.

Further evaluation plans

The major problem which needs to be solved is whether and how the factors so far not yet sufficiently controlled can be effectively controlled in the future.

The following points, to mention only a few, need to be considered:

- If one wanted to eliminate the uncontrolled influence of the school factors which do not depend on the form of organisation, one would have to set up parallel classes in both the traditional and experimental schools which were the same in teaching objectives and methods, materials used, and in assessment procedures, in so far as differences in these respects are not an unavoidable consequence of the compared school models.

(1) A report on the non-cognitive effects is at present being prepared.

- If one were to dispense with parallelism in the experimental and control schools beyond the selection and organisation of subject matter to be taught, then this would make it easier to make rough comparisons of attainment. In this case, the effects of school organisation would not be determined separately from the effects of curricular factors. One would need to examine exhaustively whether data gathered in this manner would be usable as a basis for decisions.
- It is practically impossible to control the teacher factor by strict randomisation. The secondary modern school teachers (Hauptschule) taking part in the experiment, however, can be considered as at least an approximate random sample if (a) they have no influence on the selection of their schools as experimental or control schools; and (b) they are not teaching in the experimental classes on a volunteer basis.
It is almost unavoidable that the group of Gymnasium teachers instructing in the top attainment groups in the experimental schools deviate considerably from the essential characteristics of a random sample. Their co-operation is entirely on a voluntary basis.
- If it is to be determined how pupils who would normally attend the Gymnasium in the traditional school system would learn and behave in a comprehensive school type, "experimental islands" would have to be established in which the pupil would attend an experimental school instead of a Hauptschule or Gymnasium.
- Finally, one would have to ensure that all points of view relevant to decision-making would be taken into consideration in the evaluation of the school experiments. For this purpose, the responsible groups would have to clarify the question as to whether information about the pupils' learning progress and well-being and about the contribution of the compared school models to the realisation of equality of educational opportunity, is sufficient as a basis for decision — or whether additional information would be necessary.

PART II

LEARNING TESTS

Every comprehensive discussion on the question of measurement of progress should proceed from the fact that scholastic attainment constitutes only one of several educational objectives.

Within the framework of the development of measurements of progress, the following objectives

should be considered, in addition to scholastic attainment:

- techniques of learning and of problem-solving,
- interests and attitudes,
- self-assessment and self-esteem,
- independence in thinking, evaluating and decision-making,
- ability to experience (e.g. not only formal, historical, or sociological studies of literature, music and the fine arts, but also the ability to enjoy pleasurable experiences),
- ability to co-operate, to cope with conflicts, and to establish satisfactory personal relationships.

A scientifically-based development in the area of assessment presupposes both sufficient insight into the mutual relationships between assessment and educational objectives, and a determination of the various functions of assessment in school and society.

The measurement of scholastic progress essentially has to fulfil the following three kinds of functions:

- providing data as a basis for decisions about the future educational or vocational choices of the pupil (decision-making function);
- fostering learning (didactic function);
- furnishing basic materials for effectiveness tests and analyses of scholastic measures and facilities (evaluative function).

The traditional system of measuring progress (assigning grades) fails especially with regard to its didactic functions. It often has a detrimental influence on the learning process, and seems to influence the development of the pupil's personality negatively.

Scholastic attainment testing should serve directly to promote the learning processes of the individual pupils; and, at the same time, deliver information to the teacher for more effective instruction.

The functions which concern the pupil directly are:

- the awakening or the strengthening of the desire to attain scholastic objectives (motivating function);
- the independent, effective planning of learning activities (controlling functions);
- the starting of learning processes (actualising function);
- the development of personality characteristics, as e.g. ability to learn, confidence in success, independence, interest in the subject matter (educative functions).

An effective system of scholastic attainment testing requires special methods and test materials. These have to be integrated into the instruction in such a way that the motivating, controlling, actualising and educative functions can become fully effective.

Motivation

Attainment testing for instance, could fulfil its motivating functions more effectively by means of the following procedures:

1. Immediately after the introduction of a new topic — e.g. by a lecture on the part of the teacher, a class discussion, an experiment, group or individual work — the pupils are given the opportunity to apply what they have just learned to the solution of test problems. Subsequently, they evaluate their work themselves and determine their progress in every detail.

It may be assumed that

- many pupils, in such a situation, will pay particular attention to introductory instruction since they know that they will immediately be able to apply what they have just learned;
 - many pupils will want to remove gaps in their knowledge or understanding as soon as they become aware of them;
 - every success which becomes immediately apparent to the pupil represents an incentive for further learning;
 - in the course of time, some pupils will develop playful or "sportive" attitudes which promote a healthy motivation to learn;
 - some pupils will readily take advantage of the opportunity to practise when they know they will later need to take tests in which grades will be assigned.
2. After the tests have been corrected, the teacher asks for a show of hands to determine which problems were particularly difficult, and then uses learning activities adjusted to the respective difficulties. Subsequently, the pupils receive additional test items as homework which are relevant to the topic being studied. These items are designed to offer the pupil further opportunities to practise. Since he can immediately evaluate his work, he can determine whether or not he has reached the formerly unattained objectives.

In a subsequent period of instruction a third and final test is given on the subject matter, of

which the pupils have by then a better grasp. This test, too, is immediately corrected by the pupils.

- It may be assumed that the pupils will apply themselves with considerable interest to learning activities immediately after having recognised the difficulties in the first test. This is especially true since they know they will be able to close the gaps in their knowledge or understanding and that they will be able to apply what they have just learned as soon as they tackle their homework.
- The pupils are motivated to do their subsequent homework conscientiously if they really want to achieve the goals they have not yet reached; and if they want to see how much more they have learned since the first test.
- Following the evaluation of their homework, the pupils attempt to close the remaining gaps in their knowledge and/or understanding by thinking over or practising the tasks not yet mastered, since they will want to achieve good results in the third test scheduled for the next class period.

3. The pupils enter all test results onto a "pupil's record sheet" which, by means of a graph, shows their scholastic progress both from test to test within the learning unit, and from learning unit to learning unit. The pupil's record sheets also provide continuous information for the teacher about the individual pupil's learning progress. They are not used as a basis for grading.

The scholastic progress which can regularly be read of the pupil's record sheet can serve as an effective reward to the pupil for his efforts and should stimulate him to build long-term goals.

4. The test problems are adjusted to the pupil's abilities. In general, none will be able to complete all test items correctly the first time through the test, but all pupils are able to solve a considerable number of the items. Thus, the experiences of failure and discouragement are largely avoided. Each pupil can feel confident about his future success and satisfaction with his learning progress, even though he usually still needs to learn several things in the learning unit after his first time through the test.

The shortcomings of the traditional form of instruction stand out clearly when we consider the possibilities for formative attainment testing described above.

- The pupil is often not motivated to follow the class instruction if he does not expect an examination or a test until much later, and if he does not know whether it will deal with the particular topic being covered at the moment.
- Even classes with active participation on the part of the pupils (discussion, question and answer games) will probably not produce the same motivation as scholastic attainment testing, since these forms of instruction appeal to the individual pupil only sporadically and superficially.
- In the traditional type of instruction, the pupil is confronted by a vaguely recognized, threatening pile of learning tasks which he would be most inclined to "forget". By a well-organized system of scholastic assessment he would be confronted by a well-defined set of goals. He would also be able to determine to what extent he had already achieved them and how he could learn to master them better.
- In the traditional form of instruction, the pupil is not informed about his individual learning deficiencies promptly and systematically. As a result, he cannot develop the desire to overcome these deficiencies, and he runs the risk of gradually losing his foothold.
- In most cases, the results of the pupil's individual activities are not directly apparent to him. Immediate rewards are therefore not available to motivate his learning.
- School grades contribute very little towards the development of intrinsic motivation to learn more about a subject, since they represent essentially only a global index of success. Instead, grades develop general attitudes towards attainment, success and prestige of doubtful mental, hygienic and educational value.
- Low-achievers almost always receive only poor grades. They are predominantly negatively motivated towards school and learn even less, since they experience more failures than successes. In this way, they get into a vicious circle from which it is hard for the pupil to break out.
- The unfavourable learning climate which is so frequently observed in the second stream of the secondary modern schools (*Hauptschule*) is probably caused by this very mechanism. This could be eliminated by regular feedback to the pupils of their successes.

Control of learning activities

A system of scholastic attainment testing can contribute to the effective control of learning processes by continuously setting concrete goals for the pupil corresponding to his particular stage of learning.

- The tests are prepared on the basis of a well-grounded selection and structuring of the learning objectives. This means that the learning activities may be directed towards clear, relevant, and effectively sequenced objectives.
- The learning objectives are adjusted to the respective ability levels of the pupil in such a way that he is able to master them.
- The fact that success and failure are promptly fed back directs the learning activities of the pupil to the objectives not yet mastered.

In the traditional form of instruction, the pupil receives feedback about his successes in a more global way. The teacher either says that he has learned diligently, or that, with regard to a field of learning which is not very clearly defined, he has to learn "more". Under these conditions, it is difficult for the pupil who is behind in his learning to purposefully organise the necessary learning activities without outside coaching.

Actualisation of learning processes

The usefulness and the extent of application of formative attainment testing may be considerably larger if the tests are constructed in such a way as to serve not only as measuring instruments but also as learning material. In such a case, the measurement does not take up additional time, since the test serves both as a measuring and as an instructing device simultaneously.

If the formative tests did not also fulfil instructional functions, then the teacher would be confronted with the difficult task of having to consider large amounts of test data in planning classroom instruction. By the integration of the measuring and the instructing functions in "learning tests", the teacher can largely be relieved of routine teaching activities. Thus he gains elbow room for treating learning difficulties of individual pupils and for planning the subsequent instruction.

Practical experience already available shows that it should be possible to develop the use of the tests as an especially effective learning process. In this way, formative attainment testing can be smoothly built right into the instructional process.

The taking of a test will actualise learning processes most effectively if the series of problems contained in the test is organised around a hier-

archy of learning objectives. These, in turn, are to be adjusted to the existing knowledge and abilities of the pupil, so that he is able to "climb the ladder" of learning objectives.

These learning processes take place with the intensive participation of the pupil, who has to solve the problems not only by reproduction, direct application or trivial transfer of the material learned, but also through guided discovery. If the pupil meets up with exceptional difficulties, he receives help which enables him to continue working independently.

In contrast to the traditional instructional programme, the pupil is not confined to small steps and narrow learning paths. In addition to this, there is often feedback only after the completion of a whole train of thought rather than after each small learning step. Learning tests are integrated into the rest of the instruction and can be more readily modified by the teacher, both with regard to the mode of application and internal organisation.

The possible applications of learning tests are manifold and range from the automation of knowledge and skills to concept and principle learning, and to the training of problem-solving. The educational significance of the learning tests varies from subject to subject.

In traditional instruction, it often happens that the pupil — possibly from lack of interest, distraction, satiation, or difficulties of understanding — "turns off" from time to time. Or the pupil follows instruction only passively, i.e. he listens and retains some things, but fails to establish a connection between what he has heard and what he already knows about the same subject. This passive receiving leads to a momentary collection of unrelated fragments of knowledge rather than to the development of the cognitive structure. Only learning results integrated into the cognitive structure are usable and will be retained.

Learning tests may well be a suitable means of remedying these deficiencies in the traditional form of instruction. They promote a concentrated, assimilating type of learning. Thereby, these tests can make an effective contribution to developing the cognitive structure if they are well selected and the learning objectives are appropriately structured. With regard to the mode and extent of use of learning tests, the question arises as to what weight is to be given to an efficient pre-programming of instruction on the one hand and to the provision of sufficient opportunity for autonomous learning on the other hand. The relative educational importance of these two principles and the possible ways

of their integration into the framework of an educational system using learning tests will have to be investigated. One would have to proceed from the assumption that neither efficient pre-programming nor autonomous learning should have absolute priority. Like all things in education, these principles have to be evaluated from case to case to determine the extent to which they contribute to the realisation of the different learning objectives.

These ultimately have to be based on scientifically founded political decisions.

Development of personality characteristics

Learning tests can presumably:

- contribute towards developing the capacity for "oriented learning" and
- towards awakening interests in specific subject matter;
- encourage the pupil to think independently;
- strengthen the pupil's self-confidence.

Oriented learning is characterised by knowledge extending beyond the momentary learning task, to a comprehensive learning situation.

It involves, among other things, orientation towards:

- the degree of actual and desirable mastery of learning objectives;
- the learning objectives which are hierarchically subordinate, co-ordinate or superordinate to the learning goals of the moment;
- as well as the various stages, kinds and periodic sequences of learning and their significance.

Several necessary prerequisites for the development of oriented learning can be created with the help of learning test systems. They include: a better overview of what the pupil has already learned, gathering experience in setting learning objectives, effects of learning methods, organisation of learning activities, etc.

Instructional techniques which combine the viewpoints of formative attainment testing with those of autonomous learning seem most suitable for the full development of oriented learning. The point is that the pupils must use their own initiative and creativity as much as possible in working with the learning aids.

A good orientation towards learning and even a play-like occupation with the subject matter — as can be realised by means of formative attainment testing — involving frequent successful experience, offers a favourable prerequisite for the development of self-confidence, interest in the subject matter and independent thinking.

Federal Republic of Germany

Selected projects of evaluation of school reform

INTRODUCTION

Evaluation of school reform pilot projects in the Federal Republic of Germany is organised in various ways, in keeping with the federalistic structure of the entire educational system in the eleven Länder and guided by the educational and political principles held by the individual state governments. Instead of an analysis of the complex overall situa-

tion in the field of evaluation of school reform pilot projects, the following summaries of the work of institutions practicing evaluation are presented here. They offer a representative view of the overall situation in the Federal Republic of Germany, differing as they do with respect to their official status and the organisational structures of the institutions in question, as well as with regard to the type and extent of reform projects described.

BAVARIA

Evaluation of school reform pilot programmes at comprehensive and all-day schools

by D. KRYWALSKI,
Bayerisches Staatsministerium
für Unterricht und Kultur, Munich.

Evaluation of school reform pilot projects at comprehensive and all-day schools has been conducted centrally in Bavaria since September 1970, by the "Project Group on Evaluation of School Pilot Reforms" at the State Institute of Educational Research and Planning at the order of the Bavarian State Ministry for Education and Culture. Its job is to work out a basis on which to make decisions on the future organisation of schools in Bavaria by 1976. This means that pilot schools must be given as much freedom as possible in conceiving their reforms, freedom which must be limited neither by administration nor by "consulting" scientists, and that evaluation must be regarded as an efficiency check on specific measures taken at comprehensive and all-day schools.

The project group does not advise on the pilot reforms, because the project group, although headed by an educationalist working with four sociologists and two psychologists, does not have the expertise necessary to give schools advice on practical work. For future experimental work, however, there are plans to found a consultant group at the State Institute of School Educational Theory.

In the Freestate of Bavaria, the following pilot reforms are being conducted at comprehensive and all-day schools:

3 experiments at integrated comprehensive schools (North Munich, Treuchtlingen, Hollfeld).

1 experiment at a partially integrated comprehensive school (Schwabmuenchen).

8 experiments at multilateral comprehensive schools (Ebern, Feuchtwangen, Fuerth, Geretsried, Grafenau, Marktoberdorf, Nuremberg, Schongau).

16 experiments at all-day schools (Christophorus School in Berchtesgaden, Erlangen-Spardorf Grammar School, Garching Grammar School, Gefrees Intermediate School, Hilpoltstein Grammar School and Intermediate School, Primary School on Hochstrasse in Munich, Primary School on Thelottstrasse in Munich, North Munich Comprehensive School, Neusaess Grammar School, Niederalteich Grammar School, Nymphenburg Grammar School, Pullach Catholic Family Institute, Schwabmuenchen Comprehensive School, Treuchtlingen Comprehensive School, Untergriesbach Grammar School).

In order to estimate measures taken at comprehensive schools, in particular, control schools were selected from the tripartite school system where studies were conducted that permit comparisons between the comprehensive school and the control school. From the tripartite school system, grammar schools, intermediate schools and secondary modern schools in Noerdlingen, Krumbach, Koetzing and Gemünden, as well as individual classes at Munich grammar schools, intermediate schools and secondary modern schools were used as control groups.

The varied and balanced spectrum of different experiments serves to lend a very broad basis of experience to the results, so that the consequences for educational policy — which must be drawn in any case — are not based on coincidences, but rather on sound practical experience. Here, evaluation in the sense of the recommendations for comprehensive schools of the Educational Commission of the German Educational Council differ basically from those for all-day schools. Whereas the scope for evaluation of comprehensive schools concentrates primarily on determining the effects of measures taken specifically at comprehensive schools and contrasts these effects with comparable phenomena in the tripartite school system, the evaluation of all-day schools is concerned with case histories, discovering the attitudes of parents, students and teachers toward comprehensive school measures and with economic aspects.

The question complex for evaluations of comprehensive schools

The overall task of the project group on evaluation of school pilot reforms consists of control by scientific means of the success or failure of comprehensive school measures carried out at multilateral and integrated comprehensive schools in Bavaria. Here the objectives of comprehensive schools represent at the same time the criteria for an evaluation of measures. The evaluation is concerned mainly with three objectives:

Objective 1:

More ability-oriented school career decisions

A study is being conducted to determine whether any changes in school career decisions made by the parents of the pupils in the feeding areas of Bavarian comprehensive schools can be traced back to establishment of comprehensive schools. Particular attention is paid to the expectations of the parents with respect to comprehensive schools, previous achievement abilities and social background of the pupils.

If school career decisions made by the parents for the pupil are altered, then changes should become obvious when the pupil transfers to the fifth grade. Of course, changes also occur at other points of distribution, e.g., when transferring to the seventh grade, but these are characterised more by the achievements of the pupil than by the will of the parents, i.e., they are influenced more by the school system — perhaps assignment to certain courses on the basis of achievement — than by the parents.

The most radical changes in school career decisions by parents are to be expected in the feeding regions of comprehensive schools with an integrated guidance period, because no decision in favour of a particular type of school must be made on entry into the fifth grade. In the feeding areas for these schools, therefore, the desires of the parents with regard to school completion and their transfer decisions at the beginning of the fifth grade after establishment of the comprehensive school are compared with transfers made before the school was established.

In the case of multilateral comprehensive schools, parents must decide in favour either of grammar school or secondary modern school when the child enters the fifth grade. For this reason, no statistically provable changes in parents' school career decisions are to be expected.

Objective 2:

Progress in individual achievement

Learning achievement is measured with the aid of an objective test for all types of school made up by the German Institute of International Educational Research given during the guidance period in order to make a comparison of the initial positions and rates of learning increase in the individual achievement fields (vocabulary, reading comprehension, linguistic analogies, reading for information, mathematical relationships)

- between the various school systems
- between achievement courses of the integrated comprehensive schools and grammar school or intermediate school classes.

The inherent intellectual abilities of the pupils are measured with the aid of the PSB intelligence test in grades 5, 7, and 8. Initial positions are described, changes in achievements on intelligence tests from the fifth to the seventh and from the seventh to the eighth grades determined and the following comparisons made:

- the various school systems with one another
- achievements of the integrated comprehensive school with the grammar or intermediate school classes.

As there are no achievement tests for the seventh to tenth grades which may be applied to all types of school, the Bavarian State Ministry of Education and Culture has formed a working group under the direction of the head of the "Project Group on Evaluation of School Pilot Reforms" which is to determine and describe pupil achievements at

integrated comprehensive schools in German, English and mathematics (on which performance varies widely) and to test the method of evaluation.

Objective 3:

Social integration

For "social integration", the primary question is whether the comprehensive school eliminates the divisive effect of factors such as achievement, achievement differentiation, type of school and social class on informal relationships between the pupils.

Supplementary data checks

The project group on evaluation of school pilot reforms conducts studies, in addition to the three criteria-oriented studies in four fields characterised by the fact that they cannot be assigned to one single comprehensive school objective, but belong descriptively, as well as analytically, to all objectives.

- The social background of the pupils and their abilities are registered separately from all questions relevant to the investigation at all integrated, partially integrated and multilateral comprehensive schools and control schools. The various intelligence data and social-class index values are obtained for all schools, grades and settled courses, and distribution patterns are made up.
- By systematically studying the organisational structures of the pilot reform schools and by describing measures taken, descriptive data are collected on the distribution of pupils to secondary modern, intermediate and grammar schools in the seventh grade resulting from differences in the form taken by the guidance period, and a check is conducted to find out whether the parents accept the recommendations made at the end of the guidance period. The subsequent school careers of the pupils are followed in order to permit conclusions as to the reliability of the recommendations. In this connection, changes of level at multilateral comprehensive schools and changes between grades in settled subjects at integrated schools are recorded, the

social structure of the courses is described and the percentages of pupils remaining with one course and switching courses are computed.

- In addition to objectives and organisation, subjective evaluations of organisational conditions are studied, i. e., descriptions are given of how pupils, parents and teachers evaluate the comprehensive school measures. To this end a teacher and parent poll was taken. The pupil poll was taken during the school year 1973/74 in the fifth, seventh and ninth grades and is meant to give information about how the pupils evaluate
 - the frequent change of classrooms (for classroom change in stable/flexible differentiation)
 - the frequent change in group composition (for stable/flexible differentiation) under the aspect of formation of friendships
 - the change of teachers when achievement course is changed (for stable differentiation).
- As part of the project "Development of a Model for Co-operation between Research and School", experience in co-operation between the experimental schools and those who carry out the scientific evaluation in Bavaria is recorded and analysed with an eye to possible improvements in the flow of information and co-operation.

Studies on evaluation at all-day schools

Evaluation of school pilot reform projects at all-day schools began with a descriptive inventory and is being continued with empirical studies. Here the parents of all-day pupils are polled on their evaluations of the scholastic situation of their children and their ideas about and expectations from the organisation of the all-day school. Parents were also requested to state their attitudes on leisuretime activities and homework to be done at school. A separate project supported by the Federal Minister of Education and Science analyses "Economic Aspects of the all-day School" and is to give information on the effect of the length of the school day on personnel and material expenses and attempts to work out "economically optimum approximations for all-day schools".

NORTH-RHINE WESTPHALIA

Evaluation research at the North-Rhine Westphalia Sixth Form College at the University of Bielefeld

by Mrs. E. PAREY,
Pädagogische Hochschule, Hannover.

At the North-Rhine Westphalia Sixth Form College at the University of Bielefeld (OS), research, teaching and curriculum development in the sense described below are so inseparably connected that the term "Begleitforschung" (evaluation research) is insufficient to describe the research aspect of OS work. Research is not an accompaniment to the project; rather it is an integral part of it. Thus, all scientific collaborators are involved in both teaching and research, although the amount of time devoted to each varies. There is no research plan for the OS; instead there is a plan listing development projects from which research problems arise.

At the OS, thorough-going action research is conducted. This means that the traditional division of roles into research objects and research subjects is gradually being eliminated.

In order to attain this objective, training programmes are conducted which, in the context of concrete lesson plans, teach abilities necessary for their empirical testing, e. g., designing of teaching instruments and classroom observation. College students, teachers and co-workers in special research fields will in turn be the objects of studies leading to the solution of problems arising from practical teaching at the OS. The results of these studies are directly transformed into educational action.

To be able to give an outline of the planned research activities at the OS, I must first describe the scope and objectives of the OS. The OS is planned as a link between the top grammar school form and introductory university level. Its purpose is to work out conditions and solutions permitting a co-ordinated, relatively flowing, continuous transition from the general courses of Upper Secondary School to specialised basic university courses. In a period of four years, college students are led from the beginning classes of Upper Secondary School (grade eleven) to the intermediate examination in one or two university disciplines.

The OS is conceived as curriculum workshop. Its aim is to develop and test curriculum units and sequences, new methods of teaching and learning, counselling models, evaluation procedures, possible participation etc. in such a way that they can be

transferred as elements to institutions with comparable pupil and student populations.

The OS is concerned with introduction to sciences. Its aim is to develop and test university-related training courses for a broad spectrum of scientific disciplines. Introduction to science and the general learning objectives connected with this notion are thus the focus of all teaching efforts.

Realisation of these basic OS intentions is to be made possible by various forms of course organisation:

A well-grounded choice of the main subject and the vocational field is to be prepared by participation in *elective courses* (6 hrs/week per semester) in two scientific disciplines from the first year at college onwards. Opportunities and problems connected with interdisciplinary projects are to be discovered through work in a *general course* (three weeks per semester). All types of courses serve the objective of a general introductory training course, but the *supplementary course* (5 hrs per week twice a semester) does so in a special way. Here the individual student is mainly concerned with discovering the systematic context in which his elective subjects belong. Via analysis and teaching of the basic structures of scientific processes and their functions in society, students learn what science can do to solve which problems and with which means. Each semester, the student has a certain period of time in which to organise his own individual curriculum, the so-called *intensive phase* (five weeks per semester). He may pursue special individual interests and requirements and fill knowledge gaps. Two intensive phases must be used to promote his language abilities through intensive language courses. Required *practical courses* (three four-week courses during the first three years) are supposed to emphasise and insure the practical orientation of training received at the OS. Students who have not concluded an apprenticeship or have not pursued an occupation for any length of time must attend these practical courses. They are prepared and evaluated in classroom situations in which students with occupational experience also participate.

The OS will open in September 1974. 220 students will study at the OS during the first semester. Thirty-five teachers will conduct courses, plan lessons and conduct research in the context of their courses during the first year. Two co-workers in special fields of classroom research will initially work almost exclusively on research projects and conduct training programmes in research techniques. Later, however, these co-workers will take over more teaching duties and use only part of their working time for research, when a larger proportion of teachers and students invest more time in research. The OS is to be regarded as a field for further training of teachers and researchers in which they learn to recognise which question complexes are important in the college context and require specific study.

If we consider the scope of the OS and the plans for its realisation in the light of research method, we discover that the emphasis of research activities at the OS is placed on formative and summative evaluation. The formative evaluation of the project will naturally be in the foreground for the first few years. Formulations for learning objectives will be analysed and defined. Lesson planning and methods for testing achievement will be related to teaching objectives and evaluated as to their suitability. Conversion of plans to courses and the effect thereof on students and teachers will be observed (self-observation and observation of others) and determined with structured and non-structured instruments. The results of the observations and measurements will be processed and rapidly signalled back so the lesson plans can be modified if necessary. Simultaneously, methods and instruments for evaluation will be tested for suitability and further development.

Supplementary courses in particular will be the subject of detailed analyses during the first year.

RUHR TEACHERS' COLLEGE

Research projects of the Working Group for School Development Research (AFS) at Ruhr Teachers' College

This type of course makes considerable demands on both teachers and students, because there exist neither tested contents nor teaching concepts, which clearly shows its perimental character. Conduct and results of instruction in all other types of course will be documented as far as the available capacity makes this possible. Thus, there will result, among other things, possible points of departure for later specific evaluation activities.

Formative evaluation is aimed not only at course content and organisation. In accordance with the intention of the OS, counselling, acceptance and application procedures, as well as forms of co-operative work and self-administration will be modified on the basis of empirical data so that optimum forms may be found in accordance with the OS objectives.

Proof of transferability of OS products will initially be given in constant discussions with representatives of comparable institutes and publications on the situation in curriculum development. This process of mutual exchanges can lead to alterations in the expectations of the user and modifications of the supplied product.

After the first group of students has completed the course a further contribution to proving transferability of OS products will be made with the help of summative evaluation. But proof will also be given of the more satisfactory nature of OS school and college careers as opposed to traditional careers. At first only preliminary work can be done on this subject. Initial data will be collected from all accepted students, data which will be related to the OS objectives. On the basis of these data, information can be distilled at the end of the course about changes among the students with respect to these objectives. At the present time, however, no precise data can be given about the form and content of the summative evaluation.

by Professor H.-G. ROLFF,
Arbeitsstelle für Schulentwicklungsforschung,
Dortmund.

The Structural Plan of the German Educational Council and the General Educational Plan of the Federal and State Commission on Educational

Planning (BLK) call for extensive reform measures for the present German tripartite school system. The age level school (Stufenschule) is considered a

suitable classification system for the future school system. For the majority of members of the BLK, the objective is the realisation of the integrated comprehensive school within the framework of the age level school, whereas some of the states are still reserving a decision as to type and extent of co-operation and integration in the future school system. Basically, however, the demand still exists that problems or re-organisation be made the object of scientific research and development; only in this way can sound solutions to practical problems in school reform be formulated.

This re-organisation problem is the subject of research at the Working Group for School Development Research (AFS). It is being treated in various projects, each under different aspects. In each case, the objective is to develop models and planning guidelines through interdisciplinary research. These models and guidelines should help those concerned to master problems arising in connection with re-organisation at all levels of action. Such models can only fulfil their purpose, however, when they are connected with the development of new curricula and suitable forms of social organisation in schools.

For this reason the AFS begins with a complex definition of school development planning (SEP). It defines SEP not only as "physical planning", i. e. planning of school infrastructures to meet requirements, or rather demographic forecast and corresponding scheduling and location of new, supplementary and re-organised equipment. Rather, the AFS wants to conduct its research and development work with an eye to "substantial" school development, i. e., to co-ordinate it with the development of new curricula and corresponding teaching and learning forms. In principle, such a complex or extensive definition of school development planning affects all levels of school administration: the individual school, comprehensive school supervisory boards, communities and state ministries, training and research.

School development planning can only do justice to the interests of those involved and only be realised in a democratic manner, however, when it is conceived as "school-related". For the AFS, school-related means that work is not done on development of isolated experimental schools, but rather that transferable models are formulated in close co-operation with school and administrative practice. School-related also means that starting points of participation must always be taken into account. Nevertheless, it is necessary — if only in order not to be overrun by narrow-minded individual interests, — to integrate school development planning into the entire social and educational framework. Thus it is absolutely necessary to

strive for systematic connections between individual development models and social change in existing schools.

This definition of complex school development planning, here outlined only briefly, will be explained more fully in the following sketch, which attempts to present the individual components and interdependencies of school development.

Research and development activities of the AFS are divided into four projects, which may be briefly sketched as follows:

Project A: Adaptation planning for school level system: "Seminars on School Development Planning and Development of Standardised Procedures"

Project Head: Dr. K. Klamm

In the first phase of this project, in progress since the summer of 1973, basic problems of method in school development planning (SEP) were worked on, studied and operationalised for planning of the age level school. This work was tested in training seminars for school administrative experts. Experience gathered from them has been documented in a publication on SEP methods.

Independently of the further development and especially the evaluation of the initial attempts to work out methods and processes in SEP, the work of the project also involves problems arising with the schools in realising an SEP plan after it has been set up.

Problems of implementation after setting up an SEP are formulated and dealt with primarily in the light of the following:

- Opportunities for participation in the preparation, discussion and implementation phase of an SEP
- Combining the SEP with school planning (sequence models for realisation of the planned age level school)
- Simulation models for pupil movements and evolution of pupil-numbers with alternative assumptions on transition and success quotas; comparison and refinement of prognosis techniques.

Project B: Adaptation planning for school level system: "Development of Models for Standard Equipment and School Management Systems for Lower and Upper Secondary School"

Project Head: Prof. H. Fromberger

This project is meant to study what standard equipment systems for age level schools must be like to correspond to the modified requirements of curriculum and social organisation.

The first focus of the project is:

- development of generalisable and transferable evaluation procedures for existing equipment, and,
- preparation of a catalogue of such equipment systems with the object of determining maximum and minimum equipment programmes for level schools.

The second focus of the project is concerned with improvements in school management systems in the Upper and Lower Secondary school systems, frequently larger than present ones; research subjects here are:

- to what extent proposals on reform of educational administration can be realised; and
- to what extent management techniques from administration and business can be applied to schools.

In order to answer these questions, experiences in existing school centres will be analysed.

Project C: Adaptation planning for school level system: "Development of Models for Regional and Local Advancement Planning for School Centres at Upper Secondary School Level"

Project Head: Dr. G. Hansen

Models for school centres at Upper Secondary level (S II) for local and regional advancement planning are to be developed. Using concrete examples of school pilot reforms from regions of varied structure, the pre-requisites for co-operation and integration at Upper Secondary School level are being studied.

The following questions are treated in detail:

- Location of S II centres as an integral part of regional development planning in connection with social and locational disparities;
- Definition of optimum operational sizes with

relation to curricular objectives, social organisation, etc.;

- Possible use for existing older buildings;
- Expanded use of the S II centres through extra-curricular public work (adult-education activities, youth work, etc.);
- Organisational and subject requirements of co-operation and integration.

This project is related to pilot projects in North-Rhine Westphalia.

Project D: "Social Organisation of School Centres at Lower Secondary School Level"

Project Head: K.-J. Tillmann

This project considers re-organisational problems of existing vertical school forms from the point of view of optimum internal structure of the individual school and its social environment. It is from this angle that the question is to be studied as to what conditions are necessary to permit conversion of various existing school forms in a school centre (secondary modern school, intermediate school, grammar school) into an integrated system. This project, as are the others as well, is a combined research and development project: social processes in multilateral school centres are to be studied in order to provide guidelines for development of integration models. The action proposal to be derived from this may be defined on two levels by

- indications of social environments and internal school points of departure which have a particularly positive or negative effect on integration processes; and
- indications of possible successful measures which could be taken at the various political levels (ministries of culture, school administrations, participants) to promote at least medium-range integration.

This project is related to pilot projects in Hessen.

BADEN-WUERTTEMBERG

Evaluation of school reform pilot projects in Baden-Wuerttemberg — IBS concept and initial experience

by R. H. WEISS,
Institut für Bildungsforschung und Studienberatung, Stuttgart.

General problems

The Institute for Educational Planning und Information in Stuttgart (IBS) was requested in May

1970 by the Baden-Wuerttemberg Ministry of Culture to prepare a concept of evaluation of school reform pilot projects which could be used in practice, to co-ordinate all secondary measures

and to prepare a central evaluation and analysis of the results of the experiments for the entire pilot project. A year and a half earlier, educational scientists at the Ministry of Culture were occupied under research orders with similar or other relevant questions:

Prof. G. Eigler, with the aid of a team of co-workers at Mannheim University, prepared a concept of evaluation founded on scientific theory which has now been published in Series A of the Ministry of Culture, Vol. 23.

Prof. A. Flitner (Tuebingen University) was concerned with problems of curricular re-orientation and the development of a new curriculum, particularly during the planning phase of pilot schools as part of the "Working Group Pilot Reform Schools".

Optimum "evaluation" should, in our opinion, include two essential aspects:

- evaluation as an empirical efficiency check on learning objectives, learning content and learning organisation, and
- evaluation as counselling for pilot schools.

Since each model needs a certain "starting-up" phase, the focus of activities should initially be the counselling field. If only for systematic reasons in an experimental set-up, Part I will first of all deal with the empirical efficiency check.

PART I

Evaluation as empirical efficiency check

I. Comparison of IBS concept and Eigler's beginnings

The "First Concept for Evaluation of School Reform Pilot Projects" prepared at the IBS in Stuttgart in 1970 served as a basis for measures of evaluation and counselling by the IBS. This concept, which was limited to pilot schools in Weinheim, Freiburg-Haslach, Markdorf and Bodnegg, has now been expanded to include other pilot schools: Waldkirch High School, Second Ulm Model, Heidenheim Practical High School and Osterburken All-day High School were included in the programme in 1971 at the order of the Ministry of Culture.

Prof. Eigler's report, initiated by the Ministry of Culture, proceeds from a theoretical assumption about the relationship between planning, political decision-making and empirical checks, as well as between organisational form and performance of model schools; the socio-economic aspect of evaluation, as well as the aspects of "optimum achieve-

ment advancement, social equality and social integration" are particular points of emphasis. The IBS concept, on the other hand, proceeds more from an operationalisation of the aims formulated in the individual planning groups and commissions at pilot schools, which lends a more individual and socio-psychological investigation procedures in assessing abilities, school achievement, personality dimensions and social fields. Abilities and limitations of "efficiency comparisons" between various forms of pilot schools and school systems are described in the IBS paper, in which, as opposed to the Eigler concept, no comparison is denied between the conventional school system and new pilot school systems when the absolute pre-requisite of exact definition and operational representation of general learning objectives is given at the end of Secondary School Level. The significance of the question of "Economic feasibility", to which Eigler devotes much attention, is not impaired.

II. Expanded study hypothesis

1. Pre-requisites for experiment

Under the assumption that a specific new form of school organisation is already the only alternative to existing regular schools, one could limit evaluation solely to this school form, i.e., perceive only system-related or school-related tasks (evaluation aspect). This pre-requisite, however, is not given in the case of the Baden-Wuerttemberg project programme, as it has a multi-level design. Thus an evaluating scientist trying to be objective is forced in conceiving his research design to use not only school-related or system-related hypotheses, but to study the given organisational forms of pilot schools with respect to their pedagogical, psychological, socio-cultural and economic effects comparatively as well. A "scientist" with a one-sided political and ideological orientation would relate to the one type of school reform which best conforms to his idea of a school. This is true of both main branches. The head of a reform, for example, who is oriented to the traditional Humboldt understanding of education and is, in addition, interested only in high-school "elite formation", would concentrate his studies on testing structural changes in the high-school field. It would be just as one-sided, however, for politically interested scientists to accept the general school as a matter of course already, categorically reject any other school form and thus discriminate against comparative school studies.

Methodological objections would be more to the point here; the less agreement found on common learning objectives and the more often objectives

are modified in the course of one experiment, the less decisive the results of a comparative study will be. Socio-integrative behaviour as an objective and elite consciousness are, in my opinion, mutually exclusive; they are incompatible. One of the more important conditions for comparative school research is thus agreement on and precise definition of the general learning objectives (e.g., achievement objectives in specific subjects, social integration, ability to criticise, etc). Another condition is that any school form under discussion requires counselling by trained experts. This counselling must not be limited to pupils and parents, but must also include teachers (through teacher training) and thus the classroom as well (teaching aids, curriculum development).

2. Operational definition of objectives

Basically, it must be mentioned that statements about the "value" of various school forms are just as limited and relative as statements about the "value" of a human being. The conditions of a verification of falsification hypothesis safeguarded by statistical probabilities depend on the "value references" of the investigator. If a diagnostician analyses only intellectual achievement in a single psychological study, this evaluation is certainly just as one-sided as a check on a school experiment which measures only the increase in subject knowledge. This type of study would be legitimate only if the achievement principle is the first in the hierarchy of values.

Now, however, objectives have been developed in the planning studies on all pilot schools to be evaluated; and these objectives extended far beyond the achievement principle.

A quantification of these objectives (in the form of a comparative check) is more severely limited, because as empirical studies have shown, achievement diagnoses can be made up to 90 % accuracy, but diagnoses of various personality fields (characteriological dimensions of a non-cognitive character), however, with only 50 % accuracy.

Summarised, the objectives of the pilot schools are concerned mainly with improvements (increases) in the following fields:

- improvement in cognitive function (e.g., increase in intellectual achievement or increase in learning ability; building complex linguistic structures;
- personality improvement (e.g., conflict reduction, motivation increase, overcoming fears);

- Socio-integrative improvement (e.g. overcoming social tension, social communication abilities).

The scope of problems for scientific evaluation of the efficiency of school reform pilot projects with such comprehensive objectives can thus be formulated as follows:

- a) How can I measure whether the objectives have been attained at a pilot school (school-related comparison)?
- b) How can I measure whether these objectives were attained more efficiently in one pilot school than in another comparable pilot school (inter-model school comparison)?
- c) How can I measure whether these objectives were attained more satisfactorily than in the conventional school system (inter-system comparison)?

PART II

Evaluation as counselling for school reform pilot projects

Counselling of pilot schools is an integral part of evaluation.

The following fields are part of scientific counselling of pilot schools by the IBS:

1. Pupil, teacher, and parent counselling and teaching aid

Work in this area belongs to the fields of educational counselling centres and psychologists at pilot schools. The work of a Master of Psychology at a pilot school can be summarised as follows:

— Pupil counselling:

Diagnosis, counselling and caring for development-disturbed children with learning and conflict problems; advancement diagnosis and counselling (e.g., for legasthenics).

— Parent counselling:

Counselling in case of childrens' difficulties at school and disturbed behavioural patterns; information on possible school careers and occupations for pupils, as well as on details of pilot school evaluation.

— Evaluation and counselling:

Co-operation in evaluative studies (feeding areas initial and progressive studies); participation in general teacher conferences and specialised teacher conferences on curriculum planning and

revision; counselling on improvement of achievement assessment in problems of teaching differentiation and other pedagogical and didactic measures.

2. Counselling and co-operation in curriculum development at school reform pilot projects

"Curriculum" is defined here in the broadest sense, including questions of achievement assessment and differentiation.

The curricular development work of the pilot schools should be organised on the principle of "scientific suburbs", meaning that one authority, e.g., a studies seminar, a teachers' college or a university institute, would be in charge of curricular work in one subject and that the pilot schools would be given the teaching units, tests etc. developed there. The work of the commissions at the individual pilot schools would be co-ordinated by the scientific suburb. In this system, the IBS would perform the function of documentation and transmission of this work.

GLOSSARY

Comprehensive school

"Multilateral" comprehensive school

"Integrated" comprehensive school

Grammar school

Intermediate school

(in the narrow sense)

Secondary modern school

6th form college

Comprehensive 6th form college
(academic and vocational)

All-day school

This organisational principle has not been rejected completely but nonetheless, strongly modified in practice. The commissions operate independently of one another for the most part, which is due on the one hand to practical difficulties in co-operation (lack of time, travel expenses), and on the other to special conditions at individual schools which would not permit immediate adoption of curricula developed for other schools. Since the interest of the individual commissions in curricula developed at other locations has now grown considerably, however, (individual teachers overloaded), one authority must organise documentation of all curricular work and its distribution. This function is assumed by the IBS. Useful work is only possible, however, if the pilot schools regularly send their working papers to the IBS. With some schools, this process is already being carried out to perfection, with others there are still some difficulties.

In addition to this, however, the IBS is already able to co-operate as counsellor in the individual commissions and also develop its own curricular sections and will be able to do so even more in the future.

Gesamtschule

Kooperative Gesamtschule

Integrierte Gesamtschule

Gymnasium

Realschule

Hauptschule

Oberstufenkolleg

Kollegstufe

Ganztagsschule

Norway

by Mr. O. LINDAL,

Lecturer, College of Education, Stavanger.

Educational innovation in Norway

The 1970s are from many points of view an interesting period in Norwegian school history. The reform of the primary school (the 9-year compulsory comprehensive school) has just ended with the

presentation of the new Curriculum Plan (1974), but many problems and questions are still unsolved. The upper secondary school reform has been under preparation for many years, and this year (1974) a new law introducing an all-round school system for further education (16-19) has just been pre-

sented, together with a new Curriculum Plan for this type of school. A new law, new regulations and a new Curriculum Plan for teacher training will be introduced next year (1975). The first experimental 3-year period for the district colleges has just finished.

Decentralisation of decision-making authority

Trends towards greater *autonomy* at lower levels in the educational system are present to some degree. In higher education a gradual transfer of the decision-making authority has taken place, especially where the distribution of resources is concerned, from the state to the large institutions. This is most marked at the new district colleges, where the schools' own managing bodies have comprehensive powers over the allocation of the resources at their disposal.

Like the universities, the district colleges also have great freedom in deciding the content and planning of the instruction are concerned and in their methods of evaluating the students.

At lower school levels there has been a transfer of the decision-making authority to the counties for the upper secondary school and to the communities in connection with the 9-year compulsory comprehensive schools. Work is also being done on a reorganisation of the subsidies system for the compulsory school, which will render superfluous the detailed regulations aimed at limiting the use of teacher resources. In this way important decisions on internal resource allocation can be delegated to the local authorities. The distribution of the decision-making authority between community bodies and the individual schools has not yet been clarified in this context, however.

The students' range of choices in the courses of study and subject combinations has increased at all of the secondary school levels, and the Curriculum Plan of 1974 for the 9-year comprehensive school has more of the characteristics of an indicative plan than was previously true, at the same time as it allows somewhat greater freedom in the distribution of instruction hours among the various subjects. Reduced emphasis on examinations in the 9-year school should also result in a wider range of choices where the instruction plan is concerned.

Among the vocational schools in particular there are individual examples of close contact between the school and the local community, first and foremost with local business. The new act on secondary school allows for stronger local representation more strongly into the management of the school's activities. The district colleges have, in

part, been very successful in achieving an active co-operation within their regions. In the 9-year comprehensive schools the establishment of new co-operative bodies offers possibilities for expanded contact with parents.

Until now the developments in this country towards a closer integration of the school units with the local environment have been very modest. These are, in general, cautious efforts to expand the contact surface with the local milieu, but these efforts have not really aimed at making any far-reaching changes in the decision-making process itself.

In the Norwegian educational system too, in recent years, a number of features in the decision-making process have emerged that deviate from the patterns we recognised earlier. Those trends that are being expressed can, however, in many cases be considered as conflicting trends. Increased interest in global problems goes hand in hand with greater concern with the problems of the local community. Attempts at subject integration are being made while at the same time the new, programmed instruction plans are based strictly on the "inner logic" of the traditional disciplines. Greater emphasis is being placed on the relevance to the students of the instruction, while at the same time the instruction is being programmed on the basis of given objectives, etc.

Anyone seeking a clear development trend in this area will find it difficult to find any answer in today's situation in the Norwegian educational system. Perhaps it will gradually prove that some of the seemingly conflicting trends in today's situation are not necessarily in conflict with each other at all.

The background for the reforms

The social and cultural patterns of Norway are to a large extent shaped by the geographical conditions of the country and cannot be fully understood unless seen against this background. Norway has a population of only 4 million, but in terms of area it is a fairly large country by European standards: more than 125,000 square miles (about the same size as the United Kingdom). Thus Norway is easily the most sparsely populated country in Europe with about 30 inhabitants per square mile. The isolation this creates is increased by the topography of the country and by the climatic conditions in winter. It follows from these facts that the country consists of small and in many cases rather isolated administrative units. Educational administrative authorities have to cope with a large number of small school districts and small schools.

Urban and rural schools in Norway have always operated under different conditions. Urban and rural primary schools therefore used to follow different curricula, though the subjects were the same and the pupils were supposed to reach the same standard. Up to 1959 there were two separate laws on primary education, one for schools in towns and one for schools in the country. The law of 1959 strengthened the rural school and formed the basis for further equalisation.

In secondary schools as well the differences between urban areas and sparsely populated districts is strongly felt. In Oslo, for instance, almost every student goes to a secondary school when he has finished his primary education, but this is far from the case in, for instance, Finmark.

One of the most striking features of the development of Norwegian public education is the growing wish to give all children an *equal opportunity for education*.

Three sets of objectives, with differing content, can be brought out in this context:

- Equality of access to education, without formal discrimination in relation to any group.
- Equality in the distribution of educational resources among different groups, with compensation for those groups that have a poor starting point.
- Equality in performances within the school system among different groups.

This development is closely related to general social, cultural and economic trends, and reflects the growth of *democratic forces* since the middle of the 19th century.

The educational reforms in Norway ought not to be studied in isolation; they must be compared with similar reforms in the other Nordic countries and other European countries. These societies are going through a similar development as Norway: technical innovations and increased industrialisation which create a complicated society demanding new knowledge and new abilities from its citizens. As a consequence, the school system must be in constant development in order to prepare the students for a changing society.

The education reform programme

The point of departure for the Norwegian education reform programme is to be found in the so-called Joint Programme of the Norwegian political parties announced in 1945. This includes the following statement on educational development:

"The entire school system must be co-ordinated so that the transition through its individual links, from the elementary to the highest educational level, might evolve smoothly from one to another, and this should apply equally to practical as well as literary types of schooling."

On the basis of this programme the Norwegian Labour Party, as the party in office for the next two years, assumed the task of putting into effect this programme for the *democratisation of the educational system*.

The reform of the educational system demanded an experimentation in school hitherto unknown. The experimental activity was included in a number of measures designed to improve the schools in general. These had two objects: the improvement of the "old" school and also the outlining of the new school system for the future.

Research and experiments

In 1954 it was resolved to establish a body (the National Council for Innovation in Education — NCIE) that would be of help to the Ministry "with advice, initiative and supervision on experimental activities in the school". In this context it was stated in the *Storting (the Parliament)*: "By employing *empirical science* in order to try out new school forms and educational methods, the educational system will have guided the work of reform in this important section of the social system into modern forms. These methods are being used today in every other field in our society, and they ought to be used as a matter of course in education. By utilising *running research projects* in schools, one should be able, as far as both organisation and teaching are concerned, to manage to keep education abreast of current development which is taking place in our society. It must be admitted that our school reforms have not always been made early enough to satisfy those demands which society makes on the school at any given time. This is due to a great degree to the long *study and reporting* process which has been followed up until now."

The Innovation Act gave the Ministry the following powers and responsibilities:

"His Majesty's consent having been obtained, to disregard, for purposes concerned with experimentation, the regulations embodied in the appropriate school laws, whenever such experimentation may be deemed well founded from the educational point of view and in the interest of the school."

"In order to obtain authority for experimental activity, the NCIE shall submit to the Ministry of Education plans for experimental instruction or shall pronounce on plans put forward by individual schools and others. This shall also apply to objectives and the qualifications such instruction may confer."

"Reports shall be made to *Parliament* every year on the subject of experimental activities initiated under this law."

This act might be looked upon as an interesting strategy in coping with a traditional educational bureaucracy dependent on laws and regulations. One important aspect of this law is that it gives *research and development* a change to influence entirely new educational organisations. Another aspect is its relationship to the political and administrative decision-making bodies. The Ministry has to approve the NCIE's budget and the appointments, and in policy questions the Ministry can use the NCIE deliberately as its implementation arm for policy recommendations.

NCIE activities

The National Council for Innovation in Education (NCIE) began its activities in October 1954.

Although it was stated that experiments were desirable in all types of school, it was the *9-year comprehensive school* that became the main field of activity, up until 1970. Since 1970 the main emphasis in the Council's work has been placed on experiments with and the setting up of the *upper secondary school* — the integration of the earlier *Gymnasium* and vocational education.

Most of the *teacher training* colleges have over the last 10-year period taken part in a more or less comprehensive experimental activity. In 1972 approximately 70 % of the total offers of training at the colleges were made according to the Act of Innovation.

For 20 years now the NCIE has been the main agency for promotion of large-scale educational innovation in Norway. It is closely linked to the central educational administration and to the hierarchy of the decision-making structure. The task given to the NCIE extends to almost all types of schools in Norway and to the entire country.

Types of innovations

Types of innovations in Norway can be classified in three groups:

Major reforms

One example of this type is the introduction and setting up of the *9-year comprehensive school*. Basically, the intention was to introduce a new structure and curriculum and to plan to establish this new educational provision throughout the country, school district by school district. The reform was based on proposals from several commissions and political decisions by the Ministry and Parliament.

The main role of the NCIE in this phase of the reform work has been to find answers to specific problems defined by the Ministry or the Parliament. The first phase of an educational reform of this kind is to establish the new *structure, organisation and administration*.

Curriculum development

The second phase of the above-mentioned education reform is specific *curriculum development projects*. Most of them are in our case aimed at independent learning/individualised instruction. The projects have had a more strict planning-research-and-development cycle.

These types of innovations are mainly concerned with what is assumed to be *common needs* for all the schools in the country.

Local projects

The curriculum development projects, based on the R & D approach, could reach many schools since they were based on the notion of "packaging" expert knowledge and services, which could then be disseminated throughout the country. Yet, feasible as it appeared, this strategy has not proved to meet the specific needs of the great variety of schools in the country. A new type of innovation therefore has been developed over the last years. One of its most important objectives has been to meet the needs of the individual school. Another objective is the investigation of the possibility of giving schools and local communities more freedom and initiative. This is, in itself, a major departure from a centralised strategy for innovation.

The so-called "autonomy project", "open schools" and the establishment of regional and local development centres can be looked upon as examples of a new strategy for innovation; decentralisation of decision-making and resources.

Identification of educational needs

The identification of educational needs in Norway is traditionally done through Royal Commissions which make recommendations to the Ministry and Parliament. The NCIE tries to operationalise these objectives into feasible school practice.

The establishment of priorities through commissions is seldom guided by research. The process within NCIE, however, has over the last 5-year period been more and more related to research and development projects which are carried out in order to clarify the functioning of the system as it exists. Evaluation is in some cases done outside the NCIE on contract.

Some university research projects have also proved to be of great importance for the ongoing discussion on educational needs. An example is the identification of the needs for new experiments in the 9-year comprehensive school: the 1960s were characterised by a desire to develop and strengthen the 9-year comprehensive school. Extensive experiments were initiated in order to develop suitable plans and programmes for the school's *structure, curriculum, methods, differentiation, and evaluation*.

Already early in this period it became evident that the 9-year comprehensive school was facing far greater problems than the 7-year primary school ever had. An important problem during the period of experimentation was how to create a school which all pupils — regardless of abilities and ambitions — can find interesting and rewarding, a school which can give all children opportunities to develop their special talents and potentials.

It is, of course, an extremely difficult task to develop such a school. All the problems which this task raises, are far from solved. But much thinking and work is being done in order to develop a better school for all the children.

In this connection it might be relevant to point to some of the most important results of the research study based on interviews made by Professor Sandven, University of Oslo, in 1966 and 1970. Professor Sandven's aim with this study was to find out what the pupils themselves thought about their school situation, whether or not they liked to go to school.

— About one-third of the pupils had ambivalent attitudes towards their school situation. Seven per cent of these pupils said that they were seldom happy at school. That means that 2.3 % of the total number of pupils are seldom at ease while at school.

— The pupils from the most demanding "course streams" ⁽¹⁾ were happier at school than pupils

from less demanding "course streams". According to Professor Sandven, this may suggest that our comprehensive school is a good school for those pupils who are interested in theoretical subjects, but is less suitable for those who are more interested in practical subjects.

— When asked about how they found their actual school work, a still greater number of pupils answered that they found it uninteresting. As many as 43 % of the pupils said that they disliked their school work. There were marked differences between pupils from various course streams.

— Thirteen per cent of the pupils from the ninth grade (in 1966) answered that leaving school before the final examination might be a possible solution to their problems. Among the pupils from the less demanding course streams as many as 20 % considered leaving school. But only 2.5 % of the pupils from the more demanding course streams had considered this possibility.

— In 1970, an increased number of pupils in the ninth grade wanted to leave school at once: the percentage was now as high as 17 (26 in the less demanding course streams and 5 in the more demanding course streams).

Teachers also participated in this study.

According to the teachers, the main reasons for the problems in the 7th, 8th and 9th grades of the comprehensive school are the following:

— The school does not represent an answer to the pupils' actual needs. Their need for *emotional* security and stimulation is not catered for. The school of today stresses the pupils' *intellectual* development far too much.

— The teachers "lecture" too much while at school, therefore the pupils get few opportunities for individual and creative activity.

— The schools ought to have a greater degree of autonomy. Schools are often too big, and there are often too many pupils in each class. The pupils, especially those who are less interested in theoretical subjects, should get more information about their future opportunities, both about education and working conditions in a modern society. "Vocational guidance" is a very important subject in the school, and it ought to be strengthened.

— Parents are - in many cases - too ambitious on behalf of their children. This may lead to a situation where the children feel that they are forced to do better than they are really able to, in order to live up to their parents' expectations. Parents - as well as the school - have a tendency to value theoretical capacity higher than practical capacity.

(1) The "course streaming system" meant that pupils in the 8th and 9th grades were taught with their classes, except for the subjects Norwegian, English, German and mathematics. These subjects were divided into different levels of performance, "course streams". In Norwegian, English and mathematics there were three such levels. In German — which is an elective subject — there were two levels.

Professor Sandven has given some comments upon the results of his study. He claims that the following conclusions could be drawn:

— The aim of education should be to develop all the child's possibilities and talents, not only his intellectual capacity.

— The educators' attitude towards their task ought to be changed. They ought to be more interested in the child as a growing personality than in the subject they teach. Curriculum plans ought to be more flexible. The individual school or community must be free to develop their own programmes.

— One important aim in education must be stressed: the pupils must regain their self-confidence which they - as small children - had in their own capacity to learn. They must be motivated to develop their own personality.

— The watertight division between "theoretical subjects" and "practical subjects" should be torn down both in school and in work.

— All pupils must develop the feeling of having succeeded at school. Therefore, the school must not make demands that only a small number of pupils are able to fulfil.

These conclusions reflect much of the thinking behind the third type of innovation mentioned above.

Planning of innovation

Most of the planning of innovation in Norway has until recently been done in the NCIE, which has had considerable authority in this respect. The planning has been based on wide contacts and co-operation with local authorities and teachers.

In its first years of operation the NCIE to a large extent decided the regulations to be applied for the implementation of a reform. During the 1960s one can observe a change in the role of the NCIE. In the early 1970s the planning has taken the form of consultancy and advice to the different participating bodies in an innovation. The responsibility for innovation in education is more and more divided between several agencies. Each of the new laws mentioned above (para. 1) contain one section about innovation and further development in the schools. Now the teachers, the individual schools, the local schoolboards and the different councils under the Ministry can all take the initiative, plan and carry out innovation within the broad framework of the new laws and regulations. Only the innovations that go beyond these new extended boundaries have to be carried out according to the

Act of Innovation, under the guidance of the NCIE, and in co-operation with any other council involved.

The colleges of education, for example, now have to fulfil a twofold task in innovation:

- Innovation aimed at the teacher training itself.
- Innovation in the different types of schools they train teachers for.

Model for development in structural reforms

The model for development in structural reforms has not followed a strict planning-research-development cycle.

The model is much more development-oriented. Most of the development work is done in local communities or in individual schools, based on close co-operation with the participating schools and local communities.

Evaluation in a more strict sense was not very much applied in the 9-year school reform except for evaluation of pupil performance (examinations and standardised tests).

Moving on to the secondary school reform research workers have been following the reform all the time, and there is a relatively rich crop of reports and research material on different aspects of the reform. It is fair to say, however, that revisions and new decisions are taken after various forms of assessment.

Informal and formal discussions together with the more research-oriented follow-up activities are the basis for evaluation of the secondary school reform as well as in the reform of teacher training.

In the curriculum development projects, however, there is a process which to a large extent follows the planning-research-development model: after objectives have been carefully outlined and discussed with different interest groups, a proposal is made about a curriculum development project. A project group, or a "reference group" together with a project director in the council's staff, works out prototype curriculum units and tries them out in a few schools. An evaluation and revisions cycle is based on the first or first two years of experiments, and a second version of the material is prepared and tried out in a larger number of schools with teachers who have not helped develop the material themselves. Evaluation is done by research workers from the NCIE staff or a research institute together with the project director and the project group. Based on the evaluation, a revision is made, and sometimes a third and large-scale try-out is organised. The summative evaluation also includes recommendations to the Ministry

about the implementation of the new curriculum and material in all schools in the country. Implementation is the responsibility of agencies other than NCIE.

An example: "The Mother-Tongue Project"

The background for this project is to be found in the desire to individualise teaching within the framework of the class. The main goal of the project is to organise the teaching of Norwegian in the 7th, 8th and 9th grades of the 9-year school in non-streamed classes. Developing properly differentiated teaching material is another important aspect of the project. The project is supervised by a committee appointed by the Council, and the daily leader of the project is a consultant in the Council.

The project started very modestly in 1964 with only one school. At the present stage (1974) 271 schools all over the country take part in the project. The project goes through two phases: in the primary stage preliminary experimentation in a few schools is carried out. The project is tried out, revised and tried again. In the expanded phase the project is extended to a larger number of schools, still under follow-up and supervision. Researchers outside the Council — at the University of Trondheim — study both social aspects and the professional value of the project. The expanded experiment continues for two years in each grade.

In the present project the time-table is as follows:

	1971/72	1972/73
7th grade	expand.	expand.
8th grade	primary	expand.
9th grade	primary	primary
	1973/74	1974/75
7th grade	implementation	
8th grade	expand.	implementation
9th grade	expand.	expand.

As soon as the expanded experiment has been evaluated, the teaching material that has been tried out will be finally revised and then published commercially in co-operation with one of the Oslo publishing houses (implementation).

One main problem in connection with this project has been to build up the material in such a way that it has something meaningful to offer to all categories of students. To solve this problem all texts are built up around common introductions to each chapter to be taken in full class. All questions and exercises that are connected with the introductions are differentiated. Different methods may

be used in working with these exercises. Some of the problems may be solved by groups, some of them may be solved individually, and some of them are solved by the whole class together.

Through this project valuable experience is gained in co-operation between private publishing houses and the school authorities. In some cases these two parties may have somewhat differing interests, but ways of communicating and reaching agreement have been found, and the work has been carried through with very little friction.

The aim of the project was to keep the class unit together. Nevertheless, until now the pupils have had to choose "levels" in the spring term in the 9th grade. For the examination the pupils are marked on different scales according to the "level" they have chosen. This is of course very unfortunate, but the new curriculum plan (1974) is putting an end to this.

The material produced and presented in this project has been received by teachers as well as by pupils, and representatives from these groups have been involved in material development and evaluation of the project at all stages in the process. The project as a whole has met with very little resistance. The smaller schools especially have welcomed the opportunity to keep the class unit together without having to organise different groups for each level. The mother tongue is a subject where communication, written or oral, is the main aspect, therefore, both teachers and pupils seem to find it natural that the whole class is studying this subject together.

The new curriculum plan will be introduced in the comprehensive school this autumn for all classes, except for the 9th grade. Streaming by levels is abolished by this plan, and common evaluation is introduced in all subjects. The material produced and the experience gained through the mother tongue project will probably be of great value in this connection.

An example of the Local projects: the so-called "autonomy project" in Sandnes

The project is in operation with assistance from the Work Research Institutes, has gained its inspiration from the experiments in co-operation in working life. An investigation is being carried out here into the school's technical and social organisation, and into the limiting conditions under which it is operated, with the effects which these have on influence, personal development and learning.

The experiments have not been started with any definite idea of what will be a "correct" solution

to the problems; changes must follow as a result of the school's own work on these questions. But it is obvious that the relationship between students and teachers and also the traditional methods in the division of subjects must be taken up for evaluation.

The function of research

In each of the three groups of innovation we have discussed in this paper (Major reforms, Curriculum development, and Local projects) research is playing a somewhat different role. The research function, however, is a growing concern in innovation in Norway. Nearly all major projects, run by NCIE or any other agency, are now planned in cooperation with research workers. One of the major problems, however, has been that a decision-

oriented research process is difficult to establish in reforms of a complex nature like the introduction of comprehensive schools.

Another problem has been that the money for most of the reform and development work has come from a different source than the money for the research work related to the same projects. This situation also reflects the fact that research work until recently in many ways has been looked upon as something different from reform, experimental and development work. This is now changing. Research work is viewed more and more as an integral part of all types of innovation, where the research worker has more of an assistant participant role with observation and interpretation as specific tasks in the ongoing process.

Editor: The Director of Education and of Cultural and Scientific Affairs
Strasbourg

Dr. A. Morstadt, 7640 Kehl am Rhein

Information Bulletin

Main themes in past issues

- 1/1971* **Progress Report of International Organisations 1969/1970**
Background Paper for the Seventh Conference of European
Ministers of Education
- 2/1971* **Seventh Conference of European Ministers of Education, 8-10 June 1971,
Brussels**
- 3/1971 **Assessment and Examinations — Council of Europe Studies**
A. D. C. Peterson ; M. Reuchlin ; W. D. Halls ; J. Capelle
- 1/1972 **The London Colloquium of Directors of Educational Research Organi-
sations**
- 2/1972 **EUDISED Project — Report of the Steering Group 1971**
- 3/1972 **Higher Education Reform and the Concept of Permanent Education**
J. Capelle ; W. Taylor ; B. E. Ingelmark ; E. A. van Trotsenburg ;
F. Edding
- 1/1973 **Educational Research Symposium on the Education of the 16 - 19 Age
Group**
- 2/1973 **Elighth Session of the Standing Conference of European Ministers of
Education, Bern, 5-7 June 1973.**
- 3/1973 **Symposium on Research and Reform in Teacher Education,
Bristol, 8-13 April 1973**
- 1/1974 **Symposium on the Soclo-culturally Handicapped, Ghent,
24th-28th September 1973**
- 2/1974 **Educational Research and Development**
- 3/1974 **Reform and Planning of Higher Education**

* Out of print